PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12P 19/34, C07H 21/04, A61K 37/70

A2

(11) International Publication Number: WO 95/34679

(43) International Publication Date: 21 December 1995 (21.12.95)

(21) International Application Number: PCT/US95/07605

(22) International Filing Date: 16 June 1995 (16.06.95)

(30) Priority Data:

9412054.0	16 June 1994 (16.06.94)	GB
9502728.0	13 February 1995 (13.02.95)	GB
9507640.2	12 April 1995 (12.04.95)	GB

(71) Applicant *(for all designated States except US)*: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutes of Health, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GONZALEZ, Frank, J. [US/US]; Unit 101, 5000 Battery Lane, Bethesda, MD 20814 (US). IDLE, Jeffrey, R. [GB/GB]; 19 Leazes Crescent, Newcastle, Upon Tyne NE1 4LW (GB).

(74) Agents: FEILER, William, S. et al.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: DEFECTS IN DRUG METABOLISM

(57) Abstract

The invention relates to genetic material, and specifically portions of DNA, for identifying the presence or absence of a mutation in the drug metabolism gene CYP2C9 and CYP2A6. Further, the invention comprises a method for determining such mutations and a kit incorporating the genetic material of the invention for performing the said methods so as to determine the presence or absence of mutations in the drug metabolizing gene CYP2C9 and CYP2A6.

DEFECTS IN DRUG METABOLISM

FIELD OF THE INVENTION

The invention relates to genetic material, specifically primers, for use in a method designed to determine the genotype of an individual; and also a kit, including the genetic material of the invention, for performing the method of the invention.

BACKGROUND OF THE INVENTION

It is well known that genetic polymorphisms in drug metabolizing genes give rise to a variety of phenotypes. This information has been used to advantage in the past for developing genetic assays that predict phenotype and thus predict an individual's ability to metabolize a given drug. The information is of particular value in determining the likely side effects and therapeutic failures of various drugs. The availability of this sort of information will result in routine phenotyping being recommended for certain categories of patients.

Drug metabolism is carried out by the cytochrome P450 family of enzymes. For example, the cytochrome P450 isozyme gene, CYP2C9 encodes a high affinity hepatic [S]-warfarin 7-hydroxylase which appears to be principally responsible for the metabolic clearance of the most potent enantiomer of warfarin. Similarly, the cytochrome P450 isozyme gene, CYP2A6, encodes a protein that metabolizes nicotine and coumarin and activates the tobacco-specific nitrosamine 4-(methyinitrosamino)-1-(3-pyridyl)-1-butanone) (NNK).

It is of note that the above gene products are also known to metabolize other substrates, for example, the CYP2C9 gene product is also known to metabolize Tolbutamide, Phenytoin, Ibuprofen, Naproxen, Tienilic acid, Diclofenac and Tetrahydrocannabinol.

It follows that genetic polymorphisms or

mutations in either of the two aforementioned genes can lead to an impairment in metabolism of at least the aforementioned drugs.

In so far as CYP2C9 is concerned, sequences reported by Yasumori *et al* (1987 *J. Biochem.* 102:1075-1082.) and Kimura *et al* (1987 *Nuc. Acids Res.* 15:10053-10054) show differences at several positions including a C to T base change that results in a Arginine/Cysteine polymorphism at amino acid 144. This polymorphism has been designated R144C.

In so far as CYP2A6 is concerned, a T to A base change at position 488 of the cDNA sequence described by Yamano *et al* (1990 *Biochemistry* 29:1322-1329) results in substitution of Leucine 160 by Histidine. Henceforth this mutant form of the gene will be designated CYP2A6v1.

The variant CYP2A6v1 encodes an enzyme that is unstable and catalytically inactive. It is found in the general population at a frequency of about 1% but does not account for all slow metabolizers of coumarin.

Since the cDNA sequence structure of CYP2C9 and CYP2A6 are known, and since it is also known to perform genetic assays to determine whether a preselected mutation is present within a given gene, it should, in theory, be possible to design assays which specifically determine whether either of the aforementioned mutations are present in each of the respective aforementioned genes.

However, we have found an extraordinarily high degree of exon homology in the cytochrome P450 genes. This has resulted in non-specific binding of assay materials and poor performance of assays. In the instance where primers have been used to hybridize to genetic material, non-specific binding of such primers has taken place, and in the further instance where primers have been used to hybridize to genetic material with a view to performing a polymerase chain reactions we have found that related genes have also been amplified, for example, CYP2A7, CYP2A12 and CYP2C8 have also been amplified.

SUMMARY OF THE INVENTION

The present invention relates to novel variant alleles in cytochrome P450 genes which express enzymes involved in the metabolism of particular drugs and/or chemical carcinogens.

One object of the present invention relates to the discovery of new mutant or variant CYP2A6 alleles wherein the human gene is characterized. A new variant allele has been found which is designated CYP2A6v2. The cDNA and genomic sequence of CYP2A6v2 is provided in the present invention. Another new gene related to CYP2A6 has been discovered and is designated CYP2A13. The cDNA and genomic sequence of CYP2A13 is provided in the present invention.

Another object of the present invention relates to the use of intron sequences to specifically identify CYP2A6 and CYP2C9 variants in a gene specific detection assay.

Another object of the present invention is to use an oligonucleotide probe, specific for regions unique to a particular CYP2 variant to screen for the presence or absence of the variant in a sample.

Yet another object of the invention is to provide genetic material, a method, and a kit which enable genotyping of the CYP2C9 and CYP2A6 gene with a view to providing phenotypic information concerning drug metabolism.

A further object of the present invention provides a method for diagnostically determining the sensitivity of a patient for specific drugs and chemical carcinogens. Such a method is widely applicable in determining the proper dosage of a drug for a patient.

Another object of the present invention provides a method of genotyping CYP2A6 and CYP2C9 and determining whether a mutation has altered the sequence of these genes and hence altered sensitivity to particular drugs and chemical carcinogens.

In accordance with the present invention a method is provided which utilizes the finding that each variant of a CYP2 gene has specific nucleotide differences as compared with the wild-type CYP2 gene. Such nucleotide changes can be utilized in a probe-hybridization assay, which is capable of specifically detecting a chosen variant and not other variants.

The present invention also provides a genotyping method for identifying the presence or absence of a mutation at codon 144 of the coding sequence of CYP2C9, or alternatively, at codon 160 of the coding sequence of CYP2A6, or alternatively, a gene conversion event involving CYP2A6 and CYP2A7 in exons 3, 6 or 8 comprising use of a portion of DNA. Such a mutation is then correlated to the sensitivity of particular drugs and chemical carcinogens.

The present invention further relates to a gene-specific bioassay which is capable of distinguishing between the CYP2 genes and identify the presence or absence of a mutation in CYP2A6 and CYP2C9 genes. Such a bioassay can diagnostically predict the sensitivity of an individual to particular drugs or chemical carcinogens. For example, the CYP2C9 variants identify a sensitivity to a commonly used anti-coagulant drug, warfarin. The CYP2A6 variants identify sensitivity to coumarin, nicotine and nitrosamines. The sensitivity to nicotine may be used to predict a predisposition to tobacco-related diseases, a propensity to smoking and adverse reactions to exposure to nicotine. Further, CYP2A6 genes are associated with the activation of nitrosamines, elevated levels of which have been correlated with many cancers.

The present invention also provides a method of genotyping the CYP2A6 and CYP2C9 genes using allele-specific amplification reaction.

In addition, a highly-specific combination genotyping bioassay has been developed to identify mutations within CYP2A6 and CYP2C9 which are linked to

sensitivity to particular drugs and chemical carcinogens. This combination bioassay comprises a gene-specific amplification reaction, an exon-specific amplification reaction and an endonuclease cleavage reaction wherein only one form, either mutant or wild-type is cleaved, producing either a single nucleic acid fragment or multiply nucleic acid fragments depending upon the presence or absence of the mutation. For example, one CYP2C9 variant, R144C, which contains a C_{472}→T mutation can be identified by an *Ava*II restriction site. CYP2A6 variants can also be identified by their corresponding mutations. CYP2A6v1 which contains a T_{488}→A mutation can be identified by a *Xcm*I restriction site. CYP2A6v2 which contains a T_{415}→A mutation can be identified by a *Dde*I restriction site.

The present invention also relates to a method for screening patients for drug sensitivity prior to their treatment with that drug, thereby alerting a physician of a drug sensitivity. In addition, the method may be used to screen patients for a predisposition to cancers related to excessive nitrosamine activation, which are associated with mutations within the CYP2A6 gene locus. Further, the method may be used to screen patients for a sensitivity to chemical carcinogens, based upon the genotype of the CYP2A6 and/or CYP2C9 alleles.

One such new allele variant, CYP2A6*v*2, has 98% nucleotide similarity and 80% amino acid similarity with the wild type CYP2A6, respectively. The present invention relates to the new CYP2A6v2 variant, the cDNA sequence and its genomic sequence wherein the alterations in sequence are within exons 3, 6 and 8, which are attributed to a gene conversion. In addition, another new gene, also involved in drug metabolism has been identified, and has been designated CYP2A13. This gene plays a similar role in drug metabolism as CYP2A6. These new gene sequences or fragments thereof are used as probes in identifying specific CYP2 variants in samples. In additions,

fragments of the new genes are used as primers in a genotyping assay.

The invention further provides isolated CYP2Av2 and CYP2A13 cDNAs for use in gene therapy and replacement protocols for individuals who are predisposed to sensitivity to needed drugs or to chemical or environmental carcinogens.

In accordance with an aspect of the present invention, there are provided primary human cells which are genetically engineered with CYP2A6v2 or CYP2A13 DNA (RNA) which encodes a therapeutic agent of interest, and the genetically engineered cells are employed as a therapeutic agent. (The term "therapeutic," as used herein, includes treatment and/or prophylaxis.)

Gene expression in an organism in accordance with the practices of this invention is regulated, inhibited and/or controlled by incorporating in or along with the genetic material of the organism non-native DNA which transcribes to produce an RNA which is complementary to and capable of binding or hybridizing to a mRNA produced by a gene located within said organism. Upon binding to or hybridization with the mRNA, the translation of the mRNA is prevented. Consequently, the protein coded for by the mRNA is not produced. In the instance where the mRNA translated product, e.g. protein, is vital to the growth of the organism or cellular material, the organism is so transformed or altered such that it becomes, at least, disabled.

Accordingly, in the practices of this invention from a genetic point of view as evidenced by gene expression, new organisms are readily produced. Further, the practices of this invention provide a powerful tool or technique for altering gene expression or organisms through gene therapy. The practices of this invention may cause the organisms to be disabled or incapable of functioning normally or may impart special properties thereto. The DNA of CYP2A6v2 or CYP2A13 employed in the

practices of this invention can be incorporated into the treated or effected organisms by direct introduction into the nucleus of a eukaryotic organism or by way of a plasmid or suitable vector containing the special DNA of this invention in the case of a procaryotic organism.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention are described by way of example only with reference to the accompanying figures wherein:

Fig. 1. Shows the sequence of exon 2, intron 2 and exon 3 of CYP2C8 and CYP2C9, cDNA sequences (from 4) are shown at the top of the page together with sequences from 6 genomic clones encompassing exon 2, intron 2 and exon 3 of CYP2C8 and CYP2C9. The position of the polymorphism at codon 144 of CYP2C9 and the PCR primers are indicated.

Fig. 2. Shows the sequence of intron 2, exon 3 and intron 3 of CYP2A6, CYP2A7 and CYP2A12. The position of the polymorphism at codon 160 in CYP2A6 and the PCR primers are indicated.

Fig. 3. Shows the detection of CYP2C9 Arg_{144} Cys polymorphism by PCR. Following amplification, samples were digested with *AvaII* and analyzed on a 1.8 % agarose gel . Lane I and lanes 3 to 6 show homozygous wild-type subjects, lane 2 a heterozygous individual and lane 7 undigested PCR product.

Fig. 4. Shows detection of CYP2A6 Leu_{160}. His polymorphism by PCR. Two parallel PCR reactions were carried out and the products analyzed on a 1 % agarose gel. Lanes 1, 3, 5 and 7 show the results of the wild-type specific assay and lanes 2, 4, 6 and 8 the results of the variant-specific assay for the same four subjects. Subjects I and 2 (lanes 1-4) are homozygous wild-type, subject 3 (lanes 5 and 6) heterozygous and subject 4 (lanes 7 and 8) homozygous for the mutation.

Fig. 5. Shows distribution of the weekly maintenance doses for warfarin in patients (n=57) homozygous for the CYP2C9 wild-type allele (open bars) and heterozygous (n=37) for the R144C mutant allele (solid bars). Arrows show the median weekly dose requirement of warfarin for each genotype.

Fig. 6. Represents 7-hydroxylation of coumarin (%) in a family genotyped for the CYP2A6 and CYP2A6v1 alleles, showing a subject homozygous for the CYP2A6v1 allele who is deficient in coumarin 7-hydroxylation.

Fig. 7. Shows the difference between the genomic and cDNA sequences for the CYP2A6 gene.

Figs. 8a and b. Shows the conversion event which leads to the CYP2A6v2 allele.

Figs. 9a through 9c. Shows the detection of CYP2A6v2 by PCR. (Fig. 9A) gene-specific amplification by PCR of the CYP2A6 gene using E3F and E3R. Lanes 1 to 4 show the 7.8 Kb band obtained from several representative human genomic DNA templates, lane 5 correspond to a negative control in the absence of template and lane 6 contains 1 Kb DNA ladder (GIBCO BRL) as six markers. (Fig. 9B) Exon-specific PCR amplification of exon 3 from the 7.8 Kb long-PCR product and restriction endonuclease pattern obtained after digestion with *Xcm*I (left) and *Dde*I (right) to detect the CYP2A6v1 and CYP2A6v2 alleles, respectively. The genotypes shown correspond to: wild type (+/+), heterozygous (+/-) and homozygous (-/-) subjects. (C) The genotyping strategy which has been developed. Exons are indicated by boxes. The position of the corresponding primer pairs are indicated by horizontal arrows. *Xcm*I and *Dde*I restriction sites generate digestion patterns for the different alleles having fragment sizes as shown.

Fig. 10. Schematic diagram depicting methodology underlying a CYP2C9 genotyping assay.

Fig. 11. CYP2A6v2 cDNA sequence.

Fig. 12. CYP2A6v2 genomic DNA sequence having

7216 base pairs.

Fig. 13. CYP2A13 cDNA sequence.

Fig. 14. CYP2A13 genomic DNA sequence having 8779 base pairs.

Fig. 15. Agarose minigel electrophoresis of PCR products. The CYP2C9 wild-type allele (Arg-144) and R144C respectively, Lanes marked "+/+" and "+/-" contain homozygous wild types and heterozygotes respectively. the right-hand lane contains a 100 bp ladder.

DETAILED DESCRIPTION OF THE INVENTION

The cytochrome P450 isozyme gene, CYP2C9 encodes a high affinity hepatic [S]-warfarin 7-hydroxylase which appears to be principally responsible for the metabolic clearance of the most potent enantiomer of warfarin along with metabolizing a number of other drugs and chemical carcinogens. Similarly, the cytochrome P450 isozyme gene, CYP2A6, encodes a protein that metabolizes nicotine, coumarin and a host of other drugs and chemical carcinogens CYP2A6 also activates the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (herein referred to as "NNK"). Many cancers have been associated with activation and/or accumulation of nitrosamines. The present invention allows detection of a predisposition to such cancers.

It is of note that the above gene products are also known to metabolize other substrates. For example, the CYP2C9 gene product is also known to metabolize Tolbutamide, Phenytoin, Ibuprofen, Imipramine, Naproxen, Tienilic acid, Diclofenac and Tetrahydrocannabinol and hence can also be used to detect sensitivities to these drugs. A list of CYP2C9 drug substrates has been documented and is incorporated herein by reference (Gonzalez & Idle 1994 *Clin. Pharmacokinet* 26:59-70). Hence, the present invention can be used to screen for sensitivities to these drugs.

In addition, CYP2C9 has been associated with the metabolism of chemical carcinogens, such as polycyclic

aromatic hydrocarbons. For example, the most ubiquitous environmental carcinogen, benz-[a]-pyrene is metabolized by CYP2C9. Benz-[a]-pyrene is found in tobacco, barbecued meats, car exhaust and generally, in polluted air. This compound, as it accumulates in the body becomes a potent DNA intercalating agent, ultimately resulting in cell transformation and the formation of tumors. The present invention provides a diagnostic method of screening individuals for their ability to metabolize and hence inactivate benz-[a]-pyrene. For example, a homozygote wild-type CYP2C9 individual would be better able to tolerate high levels of benz-[a]-pyrene than a heterozygote of the CYP2C9 allele.

Similarly, the CYP2A6 allele is associated with drug sensitivity and carcinogen metabolism. Coumarin sensitivity is directly related to the presence of a variant CYP2A6 allele, such as CYP2A6v1, CYP2A6v2 and also CYP2A13. Coumarin is a drug used in treatment of neoplastic diseases, such as lymphomas. (See <u>Martindale: The Extra Pharmacopoeia</u> 1993 Ed. Reynolds, J.E.F., The Pharmaceutical Press, London, p. 1358). Its suggested dosage is very high. Therefore, the present invention is useful in determining a patient's sensitivity to the drug in order to prescribe a proper dosage and avoid toxicity.

Another drug, Thiotepa™, is used in the treatment of a variety of neoplastic diseases, such as in treating women with breast cancer and children with brain tumors. Thiotepa is metabolized by CYP2A6 into Tepa, which is an intermediate more therapeutically potent than Thiotepa. Therefore, if a patient has a very active CYP2A6 enzyme, it is likely the patient will require lower doses of Thiotepa to provide a therapeutically effective amount. As one can see, the dosage provided to a patient is dependent upon the rate a patient is capable of metabolizing activating the drug. The present invention has identified variant alleles whose enzymatic activity is compromised. In addition, the present invention provides

a simple method of genotyping patients for Thiotepa drug sensitivity. With information concerning patient sensitivity to such drugs, the proper dosage can be provided, hence maximizing drug efficiency and minimizing drug toxicity.

Further, CYP2A6 has been associated with nicotine metabolism. In addition to being an active ingredient in tobacco, nicotine also has several clinical uses. Nicotine is used clinically to treat various neurological disorders, such as Parkinson's disease and Alzheimer's disease. In addition, nicotine is used to treat tobacco addiction. In all of these situations, it is important to know a patient's sensitivity to nicotine, since extremely sensitive patients will become violently ill upon administration of nicotine. Therefore the present invention provides a method of identifying nicotine-sensitive patients by genotyping a patient's CYP2A6 allele. The present invention also provides a convenient method for determining an individual's general predisposition to using tobacco based upon their sensitivity to nicotine.

In addition, CYP2A6 is involved in activating nitrosamines, thereby producing the potent carcinogen NNK. Increased levels of NNK have been associated with a variety of cancers, including but not limited to lung cancer, nasal-pharynx cancers, throat cancers and colon cancers. In general, elevated levels of CYP2A6 has been associated with cancers associated with exposure to nitrosamines. The present invention may detect a patient's predisposition to such cancers. The presence of a CYP2A6 gene or a variant thereof will affect the likelihood that procarcinogens present in tobacco smoke will be activated into carcinogenic nitrosamines and nitrosamine-derivatives and therefore result in the development of a cancer.

It follows that genetic polymorphisms or mutations in either of the two aforementioned genes can

lead to an impairment in metabolism of at least the aforementioned drugs and chemical carcinogens.

The present invention relates to the identification of the absence or presence of mutations in CYP2C9 and CYP2A6 and thus predict the phenotype of an individual and so predict whether and how an individual is likely to metabolize particular drugs and chemical carcinogens. For instance, the R144C mutation arising from a $C_{472} \rightarrow T$ base substitution in the CYP2C9 gene results in a reduction in warfarin metabolism. This implies that patients with this mutation receiving warfarin require a lower dose to maintain an anticoagulation target than those patients who do not have the mutation and are also receiving warfarin. Conversely, homozygous wild-types require higher doses in order to maintain an anticoagulation target.

"Mutation", as the term is used herein denotes an allelic variation of a known sequence, which alters the expressed gene product's activity. Such a variation need not completely inactivate the gene product's activity but merely alter it.

Similarly, one mutation within CYP2A6v1 arising from a $T_{488} \rightarrow A$ base change results in substitution of Leucine 160 by Histidine. Another CYP2A6 variant, CYP2A6v2, has been identified which differs from CYP2A6 in the regions of exons 3, 6 and 8. One particular mutation in CYP2A6v2, $T_{415} \rightarrow A$ mutation is useful in the assay of the present invention. These substitutions are very useful in detecting predispositions to cancers associated with tobacco and activation of nitrosamines. The normal CYP2A6 enzyme functions in the metabolism of nicotine, one of the carcinogenic compounds in tobacco.

In addition, the present invention relates to the identification of a new variant of CYP2A6 designated CYP2A6v2. The variations of CYP2A6v2 from CYP2A6 bear sequence relatedness with the corresponding exons of the CYP2A7 gene, suggesting a recent gene conversion. The cDNA

and genomic sequence for this gene is provided in the present invention. Hence, at least three different allelic variants of CYP2A6 exist and are illustrated in Figure 8. These allelic variants include CYP2A6, CYP2A6v1 and CYP2A6v2.

Further, the present invention relates to a new CYP2A gene, designated CYP2A13. This gene produces an inactive form of CYP2A6, however variants at particular positions, including amino acid positions 117, 209 and 365 produce an enzyme which may alter the enzyme's activity and hence affect drug sensitivity. These mutations in CYP2A6 are likely to result in a deficiency or impaired activity of one of the enzymes responsible, for example, for metabolizing drugs, nicotine and nitrosamines.

CYP2A13 is considered a new cytochrome P450 gene. However, since the CYP2A13 gene product has a similar function as the CYP2A6, it is discussed herein as a variant of CYP2A6. That is, assays using the specific mutated amino acid positions 117, 209 and 365 of CYP2A13 and detecting variations at those positions are indicative of CYP2A6-like variant functions.

In one embodiment, the CYP2A6v2 or CYP2A13 proteins or functional portions thereof are expressed as recombinant genes in a cell, so that the cells may be transplanted into an individual in need of gene therapy due to the predisposition to a carcinogen-associated cancer or a sensitivity to a drug. To provide gene therapy to an individual, a genetic sequence which encodes for all or part of the CYP2A6v2 or CYP2A13 ligands are inserted into vectors and introduced into host cells. Examples of vectors that may be used in gene therapy include, but are not limited to, defective retroviral, adenoviral, or other viral vectors (see, e.g., Mulligan, R.C., 1993, Science, 260:926-932). The means by which the vector carrying the gene may be introduced into the cell includes, but is not limited to, microinjection, electroporation, transduction, or transfection using DEAE-

dextran, lipofection, calcium phosphate or other procedures known to the skilled routineer (see, e.g., Sambrook et. al. (Eds.), 1989, In "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). Examples of cells into which the vector carrying the gene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type.

More specifically, there is provided a method of enhancing the therapeutic effects of blood cells, that are infused in a patient, comprising: (i) inserting into the blood cells of a patient a DNA (RNA) segment encoding CYP2A6*v*2 or CYP2A13 gene product that enhances the therapeutic effects of the blood cells; and (ii) introducing cells resulting from step (i) into the patient under conditions such that the cells resulting from step (i) "target" to a tissue site. In the alternative, as previously described the cells are not "targeted" and functions as a systemic therapeutic. The genes are inserted in such a manner that the patient's transformed blood cell will produce the agent in the patient's body. In the case of antigen-specific blood cells which are specific for an antigen present at the tissue site, the specificity of the blood cells for the antigen is not lost when the cell produces the product.

Alternatively, as hereinabove indicated, CYP2A6*v*2 or CYP2A13 DNA (RNA) may be inserted into the blood cells of a patient, in vivo, by administering such DNA (RNA) in a vehicle which targets such blood cells.

Further details regarding methods of gene therapy are provided in Anderson et al., U.S. Patent No. 5,399,343 which is herewith incorporated herein by reference.

In another embodiment, antisense CYP2A6*v*2 or CYP2A13 DNA or RNA may be used to control the expression of CYP2 gene. For example, antisense therapy may be used

to control CYP2A6's ability to activate dangerous nitrosamines by curbing its expression. Methods of producing such antisense molecules are described in U.S. Patent No. 5,190,931, which is incorporated herein by reference.

Developing a genotyping assay, which could distinguish the CYP2 genes of interest from other cytochrome P450 genes required careful engineering since these genes have a high degree of sequence homology. To overcome this problem, one embodiment of the present invention has elucidated the genomic sequence structure of CYP2C9 and CYP2A6 with a view to making, in part, intron specific primers. That is to say primers which, in part, hybridize to at least one intron, preferably an intron adjacent to an exon including the mutation of interest, in the gene to be examined. Since there is less homology between the introns of cytochrome P450 genes, it has been found that using intron specific primers, gene specific assay can be undertaken. The present invention has a further advantage of using intron specific primers in so far as the use of such primers facilitates the manufacture of an optimum length of DNA which in turn facilitates the specificity of the instant bioassay.

A "genotyping" assay as the term is used herein refers to any diagnostic or predictive test to detect the presence or absence of allelic variants of a known gene sequence at a specified gene locus. Two gene loci are of particular interest in the present invention, CYP2A6 and CYP2C9.

Further, the present invention relates to differences between the genomic DNA sequence structure and the cDNA sequence structure, as illustrated in Figure 7. As a result, primers directed at the genomic sequence structure have been developed which are more reliable.

Several methods are provided for identifying the presence or absence of a mutation at codon 144 of the coding sequence of CYP2C9, or alternatively, at codon 160

of the coding sequence of CYP2A6, or alternatively, a gene conversion event involving CYP2A6 and CYP2A7 in exons 3, 6 or 8 comprising a DNA encompassing the region of a CYP2 gene unique to that variant.

One such method relates to an assay which contemplates the use of one specific primer which specifically encompasses the region containing the mutation, and a second primer which is complementary to another portion of the gene. The second primer sequence chosen is based upon the CYP2A6, CYP2C9 or CYP2A13 sequences as set forth in figures 12, 1 and 14, respectively, depending upon the preferred size of the amplification product. One skilled in the art will know how to select second primer based on the region of gene chosen for amplification. These primers need not be identical to a given sequence but must be sufficiently complementary to hybridize to the target region in a specific manner. In short, the primers are preferably at least substantially homologous to the nucleic acid sequence provided.

Nucleic acid sequences includes, but is not limited to, DNA, RNA or cDNA. Nucleic acid sequence as used herein refers to an isolated nucleic acid sequence. Substantially homologous as used herein refers to substantial correspondence between the nucleic acid primer sequence of as described herein and that of any other nucleic acid sequence. Substantially homologous means about 50-100% homologous homology, preferably by about 70-100% homology, and most preferably about 90-100% homology between the particular sequence discussed and that of any other nucleic acid sequence.

In the instant application, the term "primer" is further used to designate a molecule comprising at least three nucleotides, the exact length being determined by the requisite amount of DNA needed, under given reaction conditions, to bind to or interact with a test sample so as to identify the presence or absence of either of said

mutations. Preferably, the primer is usually between 15 and ideally about 20 to 50 oligonucleotides in length.

The primer is selected, or adapted, to be substantially complementary to a part of DNA which is adjacent to the region of at least one of the aforementioned mutations. Thus such a primer is able is hybridize with a part of DNA that contains a region in which the mutation of interest may be found. Although the primer may not reflect the exact sequence of the region in which the mutation is thought to occur, the more closely the primer is to this sequence, then the better the binding will be. Ideally, the more closely the sequence of the 3′ end of the primer is to said region the better the binding or interaction will be.

An alternative method for using the sequence unique to a variant for detection relates to use of an oligonucleotide probe for specifically detecting the presence or absence of a CYP2 variant gene in a sample. this method comprises the steps of contacting the sample with a nucleic acid probe, allowing hybridization, forming a probe: CYP2 variant complex; washing excess probe from probe: CYP2 variant complex; and detecting probe: CYP2 variant complex, wherein a positive signal is an indication of the presence of the CYP2 variant in the sample.

The hybridization of the probe to sample nucleic acids can be carried out by any of the methods commonly used in the art. Such methods include but are not limited to, Dot blot, Colony hybridization, Southern blot, solution hybridization and *in situ* hybridization.

Washing the excess probe from the probe: CYP2 variant DNA can be accomplished by many well-known methods. Simply rinsing the complex with excess buffer will facilitate removal of excess probe. Alternatively, washing may entail separating the probe: CYP2 variant complex from excess probe. Many methods are known to one skilled in the art and include but are not limited to

centrifugation, filtration and magnetic force.

According to the present invention there is provided a portion of DNA suitable for use as a primer in a method for identifying the presence or absence of a mutation either at codon 144 of the coding sequence of the gene CYP2C9, or alternatively, at least one gene conversion event involving CYP2A6 and CYP2A7 in exons 3, 6 or 8, or alternatively, at codon 160 of the coding sequence of the gene CYP2A6; comprising a DNA which is adapted to hybridize to at least one intron of at least one of said genes.

In one embodiment, the method comprises the use of at least one restriction endonuclease to digest DNA from individuals to be tested. In this instance, DNA from individuals positive for the wild-type form of CYP2C9 provide a digest with a restriction endonuclease, such as *Ava*II results in production of two fragments, a first fragment including 270 base pairs and a second fragment including 50 base pairs. In contrast, individuals having the aforementioned mutation in CYP2C9 present a single fragment of 320 base pairs only. This is due to a loss of the *Ava*II site. The CYP2A6 gene variants can also be distinguished by the occurrence of specific restriction endonuclease sites. The CYP2A6v1 variant, which is a T_{488}→A mutation in exon 3 can be identified by a variant-specific *Xcm*I restriction site. The CYP2A6v2 variant, which contains a C_{415}→A mutation within exon 3 can be identified by a variant-specific *Dde*I restriction site. The wild-type CYP2A6 gene does not contain either an *Xcm*I or *Dde*I site. The results of such restriction endonuclease digestions are illustrated in Figure 9.

It may be necessary to amplify the DNA prior to digestion. Such may be the case when the DNA of interest is present in minute quantities in a sample. In such circumstances, amplification of DNA to be tested is undertaken before digesting the DNA as described above. This provides for a greater quantity of materials.

Amplification is performed using any conventional technique, such as by a PCR reaction. Many other techniques for amplification can be used in producing sufficient DNA for detections. Such amplification techniques are well-known to the skilled artisan and include, but are not limited to polymerase chain reaction (PCR), PCR *in situ*, ligase amplification reaction (LAR), ligase hybridization, QB bacteriophage replicase, transcription-based amplification system (TAS), genomic amplification with transcript sequencing (GAWTS) and nucleic acid sequence-based amplification (NASBA). A general review of these methods is available in Landegren, et al., *Science* 242:229-237 (1988) and Lewis, R., *Genetic Engineering News* 10:1, 54-55 (1990), which is incorporated herein by reference.

One embodiment of the present invention uses oligonucleotide primers in an amplification and detection assay. A basic description of nucleic acid amplification is described in Mullis, U.S. Patent No. 4,683,202, which is incorporated herein by reference. The amplification reaction uses a template nucleic acid contained in a sample, two primer sequences and inducing agents. The extension product of one primer when hybridized to the second primer becomes a template for the production of a complementary extension product and vice versa, and the process is repeated as often as is necessary to produce a detectable amount of the sequence.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E.coli* DNA polymerase I, thermostable *Taq* DNA polymerase, Klenow fragment of *E.coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase and other enzymes which will facilitate combination of the nucleotides in the proper manner to form amplification products.

A sample being screened for the presence or absence of a mutation in CYP2A6 and/or CYP2C9 genes can be tested with the instant invention. The nucleic acid material can be in purified or nonpurified form, provided the sample contains the CYP2A6 and/or CYP2C9 genes. The sample may be derived from any tissue or bodily fluid, wherein the patient's DNA can be found. A clinically practical type of sample is a blood specimen which contains patient DNA and can conveniently be genotyped in the bioassay of the present invention.

The "primers", as the term is used in the present invention refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions wherein synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e. in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH. The primers are preferably single stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare amplification products. Preferably, the primers are oligodeoxyribonucleotides. The primers must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers will depend on many factors, including temperature, source of primer and use of the method. For diagnostic methods, the primers typically contain at least 10 or more nucleotides. The oligonucleotide primers may be prepared using any suitable method, such as, for example, the phosphotriester and phosphodiester methods (Narang, S.A., et al., *Meth. Enzymol.* 68:90 (1979); Brown E.L., et al., *Meth. Enzymol.*, 68:109 (1979)) or automated embodiments thereof. In one such automated embodiment diethylphosphoramidites are used

as starting materials and may be synthesized as described by Beaucage et al., *Tetrahedron Letters* 22:1859-1962 (1981). One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Pat. No. 4,458,066. It is also possible to use a primer which has been isolated from a biological source (such as a restriction endonuclease digest).

In a genotyping bioassay of the present invention, one embodiment comprises a gene-specific amplification reaction, an exon-specific amplification reaction and a restriction endonuclease reaction. In such a reaction a suitable polynucleotide polymerase is used in the amplification reaction, many of which have already been described in the art. In addition, any appropriate restriction endonuclease which is designed to digest the DNA and so provide information concerning genotype may be used.

It may further be necessary to provide a label on the nucleic acid for detection. The nucleic acid can be DNA or RNA and made detectable by any of the many labeling techniques readily available and known to the skilled artisan. Such methods include, but are not limited to, radio-labelling, digoxygenin-labeling, and biotin-labeling. A well-known method of labeling DNA is ^{32}P using DNA polymerase, Klenow enzyme or polynucleotide kinase. In addition, there are known non-radioactive techniques for signal amplification including methods for attaching chemical moieties to pyrimidine and purine rings (Dale, R.N.K. et al. 1973 *Proc. Natl. Acad. Sci. USA*, 70:2238-2242; Heck, R.F. 1968 *S. Am. Chem. Soc.*, 90:5518-5523), methods which allow detection by chemiluminescence (Barton, S.K. et al. 1992 *J. Am. Chem. Soc.*, 114:8736-8740) and methods utilizing biotinylated nucleic acid probes (Johnson, T.K. et al. 1983 *Anal. Biochem.*, 133:125-131; Erickson, P.F. et al. 1982 *J. of Immunology Methods*, 51:241-249; Matthaei, F.S. et al 1986 *Anal. Biochem.*, 157:123-128) and methods which allow detection by

fluorescence using commercially available products. Non-radioactive labelling kits are also commercially available. Such a label can readily be incorporated into the nucleic acid during an amplification step. In the absence of an amplification step, a target nucleic acid can readily be chemically or enzymatically modified to carry a label. Additionally, it may be preferable to provide a labeled primer which may serve to incorporate a label into the nucleic acid target. Probes, as may be used in an embodiment of the invention may also be chemically or enzymatically labeled as described above.

In a preferred embodiment of the invention said DNA primer hybridizes to an intron adjacent said position of said mutation. Preferably said DNA is a primer with the 3'-end specific for the gene of interest. Preferably further still said DNA is single stranded. Preferably further still, in so far as the CYP2C9 mutation is concerned, said primers are as follows:

HF18: position 8 of intron 2 onwards of genomic sequence in forward orientation comprises
5' TGCAAGTGCCTGTTTCAGCA 3'
HF2R: position 505 onwards of cDNA sequence in reverse orientation comprises
5' AGCCTTGGTTTTTCTCAACTC 3'.

It is of note that both these primers are designed to be specific for CYP2C9 and so do not amplify related genes such as CYP2C8, which notably also has an Arginine_{144} present.

Preferably, in so far as CYP2A6 is concerned, three primers J51, J61 and B are used in two parallel allele-specific PCR reactions. These primers are as follows:

J51 comprises 5' GGCTTCCTCATCGACGCACT 3'
(forward strand from position 479 of cDNA

sequence described as hIIA3 (Yamano, et al. 1990 *Biochem* 29:1322-29)).

J61 comprises 5' GGCTTCCTCATCGACGCACA 3' (forward strand from position 479 of cDNA sequence described as hIIA3v (Yamano, et al. 1990 *Biochem* 29:1322-29)).

Both J51 and J61 contain a substitution at position 18 of A for C to give improved specificity as suggested by Newton *et al* (1989 *Nuc. Acids Res.* 17:2503-2516).

Primer B comprises 5' AATTCCAGGAGGCAGGGCCT 3' (reverse orientation from position 125 of intron 3 of CYP2A6 (onwards). Designed so that only CYP2A6 and not CYP2A7 or CYP2A12 are amplified.

One method of genotyping CYP2A6 provides an allele-specific amplification reaction method is used. In this instance, DNA which is adapted to specifically hybridize to the wild-type or the mutant type of the gene is incubated with test DNA under reaction conditions and the resultant products are analyzed by electrophoresis and then visualized by staining with ethidium bromide. Individuals who are homozygous for the wild-type allele produce a reaction product with primer J51 only. Similarly, individuals who are homozygous for the mutation produce a reaction product with primer J61 only. Those individuals who are heterozygous produce a reaction product with both J51 and J61.

Alternatively, another method for genotyping CYP2A6 is provided in a specific amplification bioassay, which is achieved with primers *F4* and *R4* as follows:

The F4 primer (forward) comprises 5' CCCCTTATCCTCCCTTGCTGGCTGTGTCCCAAGCTAGGCAGGATTCATGGTGGGGCA 3', wherein a preferred fragment thereof further comprises 5' CCTCCCTTGCTGGCTGTGTCCCAAGCTAGGC 3'.

The R4 primer (reverse) comprises 5′ GCCACCACGCCCCTTCCTTTCCGCCATCCTGCCCCCAGTCTTAGCTGCGCCCCTCTC 3′, wherein a preferred fragment thereof further comprises 5′ CGCCCCTTCCTTTCCGCCATCCTGCCCCCAG 3′.

This method of CYP2A6 genotyping involves a first amplification reaction with F4 and R4 primers, which generates a DNA fragment approximately 7.8 kb in size. This amplification step is facilitated by polymerases which are capable of transcribing long stretches of DNA. To distinguish the CYP26Av1 and CYP26Av2 variant alleles, an exon-specific amplification step is carried out using the 7.8 Kb DNA fragment as template DNA. This may be accomplished using the following primer pair:

The E3F primer (forward) comprises 5′ CCTGATCGACTAGGCGTGGTATTCAGCAACGGGGAGCGCGCCAAGCAGCTCCTG 3′, wherein a preferred fragment thereof further comprises 5′ GCGTGGTATTCAGCAACGGG 3′.
The E3R primer (reverse) comprises 5′ CGCGCGGGTTCCTCGTCCTGGGTGTTTTCCTTCTCCTGCCCCCGCACTCGGGATGCG 3′, wherein a preferred fragment thereof further comprises 5′ TCGTCCTGGGTGTTTTCCTTC 3′.

Using these primers in a second amplification reaction step a segment of CYP2A6 exon 3 is specifically amplified. The method further comprises use of the restriction endonuclease *Xcm*I to detect the CYP2A6v1 mutation and *Dde*I to detect the CYP2A6v2 mutation.

According to a yet further aspect of the invention there is provided a kit for performing the afore described methods which kit includes at least a portion of DNA in accordance with the invention and preferably at least one control sample of DNA containing the mutation or

mutations of interest and ideally also a wild-type sample of DNA so that suitable comparisons can be made.

It is of note that although the method is described with reference to the above methods, any suitable method using the genetic material of the invention may be used to identify the mutations described herein.

The CYP2C9 assay has been used in a study of warfarin dose requirement in 94 patients undergoing anticoagulant treatment and the results obtained are summarized in Figure 5. 58 patients (61.7%) were homozygous for the wild-type (Arg_{144}) allele and were found to require a median weekly maintenance dose of 31.5 mg of warfarin. 36 patients (38.6%) were heterozygous and required a median weekly maintenance dose of 24.5 mg. The doses required by the two groups were significantly different (Mann-Whitney U-test, $p = 0.016$). No subjects in the group were homozygous for the mutant allele but based on allele frequencies and the Hardy Weinberg equilibrium, the predicted frequency of homozygous mutant subjects is 3.7%.

Comparison of the weekly maintenance dose of warfarin in the R144C heterozygotes ($n = 36$) and homozygous wild-type ($n = 58$) reveals that the heterozygotes required a significantly lower dose (range of 10.5 - 80.mg). Moreover, of the patients requiring the lowest doses to maintain an anticoagulation target (INR 2.0-4.0), in the range 5-15 mg per week, 9 out of 10 were heterozygous. At the other extreme of weekly doses >55 mg, 5 out of 6 patients were homozygous wild-type for CYP2C9. The significantly lower (20%) warfarin dose requirement of the patients with one variant R144C allele is consistent with the kinetic properties of the R144C protein with respect to (S)-warfarin hydroxylation and presumed *in vivo* metabolic clearance (Rettie et al. 1994 *Pharmocogen.*, 4:39-42).

The CYP2A6 genotyping assay has been used in

studies on coumarin metabolism. Coumarin 7-hydroxylase activity is a convenient marker activity to identify the presence of CYP2A6 in a particular sample. There is considerable variation in the ability of individuals to 7-hydroxylate this compound which is a reaction specific for CYP2A6. A subject deficient in coumarin 7-hydroxylation has been identified. This subject is homozygous for the mutant CYP2A6v1 allele confirming the previous *in vitro* findings that substitution of Leu*160* by His results in loss of coumarin 7-hydroxylase activity. As shown in Fig. 6, CYP2A6 genotyping and phenotyping with coumarin has been performed on other members of the proband's family and impaired coumarin 7-hydroxylation has been observed in heterozygotes for the CYP2A6v1 mutation.

The genotyping assays described herein resulted from a two step amplification reaction wherein first amplification reaction amplifies a 7.8 Kb fragment containing the CYP2A6 gene (Fig. 9A) and a second amplification reaction amplifies an exon-specific fragment of CYP2A6. The amplification product was digested with restriction endonucleases producing different patterns for the various CYP2A6 alleles. Representative results obtained for several human subjects for the detection of the CYP2A6v1 (*Xcm*I digestion) and CYP2A6v2 (*Dde*I digestion) are shown in Figure 9 panel B. A schematic depiction of this genotyping assay is shown in Figure 9, panel C. Of 155 human genomic DNA samples analyzed 21 heterozygous (+/-) and 6 homozygous (-/-) subjects were detected for the CYP2A6v1 allele, whereas 17 heterozygous (+/-) and no homozygous were identified for the CYP2A6v2 allele variant. Additionally, 7 homozygous for both CPYP2A6v1 and CYP2A6v2 alleles were found.

Allelic frequencies were calculated for either allele in several ethic groups and analyzed as shown in Table 1. CYP2A6v1 frequency is almost identical between Caucasian and Japanese, and it is only twice the frequency in Taiwanese samples. Significantly, this allele is

completely absent in the African-American population within the samples studied. The Japanese population has a remarkable higher frequency for the CYP2A6v2 allele (28%) as compared to the Caucasian (2%), Taiwanese (6%) or African-American (2.5%) (ethnic groups).

Table 1: Allelic frequency for the CYP2A6 gene in different ethnic groups.

	Allelic Frequencies (%)			
Ethnic Group	CYP2A6	CYP2A6v1	CYP2A6v2	N
Caucasian	75	23	2	52
Japanese	52	20	28	40
Taiwanese	83	11	6	178
African-American	97.5	0	2.5	40

The following examples illustrate various aspects of the present invention and in no way are intended to limit the scope thereof. All books, articles, and patents referenced herein are incorporated herein, *in toto*, by reference. Other similar embodiments will be clear to the skilled artisan and are encompassed within the spirit and purview of the present invention.

EXAMPLE 1

Method for determining the genotype CYP2C9

Genotyping for the CYP2C9 polymorphism is carried out by amplification by PCR followed by digestion with the restriction endonuclease *Ava*II. Amplifications are performed in 0.5 ml microcentrifuge tubes in a volume of 100 μl containing 10 mM Tris-HCl, pH 8.8, 1.5 mM MgCl2. 50 mM KCl, 0.1% Triton X-100, 5% dimethylsulphoxide, 200 μM each of dTTP, dATP, dCTP and dGTP, 250 μM of the primers HF18 and HF2R, 2.5 units Taq polymerase and 1 μg human leukocyte genomic DNA. PCR conditions consist of 35 cycles with a denaturation at 93°C for 1 min. annealing at 55°C for 1.5 min and polymerization at 72°C for I min. 20 μl of the amplified DNA is incubated with 10 units *Ava*II for 3h at 37°C and then analyzed by electrophoresis on

1.8% agarose minigels in TBE (90 mM Tris-borate, 2 mM EDTA) buffer. The digestion products are visualized by ethidium bromide staining. DNA from individuals positive for the wild-type Arg,44 is digested to give fragments of 270 bp and 50 bp whereas in individuals with the mutant Cys_{144} present, a band of 320 bp is seen due to loss of an *Ava*II site (Figure 3).

EXAMPLE 2

Genotyping for the CYP2C9 polymorphism was carried out by amplification by PCR followed by digestion with the restriction endonuclease *Ava*II.

One hundred patients were recruited from two anticoagulation clinics in the Newcastle area over four study days. Body weight and height were measured, the basal metabolic index ("BMI") calculated for each patient and details of age, sex, drug history, current and previous International Normalized Ratio ("INR") determinations, indications for anticoagulation and other significant health problems were all recorded. DNA was isolated by a standard manual chloroform-phenol extraction procedure and 1μg was subjected to PCR analysis. As shown in Figure 10 the C→T substitution, which converts Arg→144 to Cys, resides in exon 3 of the CYP2C9 gene and results in the loss of an *Ava*II restriction site (...GAGGACCGTGTTCAA...) in the R144C allele (...GAGGACTGTGTTCAA...). This provided the basis of the amplification strategy. A CYP2C9 specific intron forward primer (HF18, TGCAAGTGCCTGTTTCAGCA, Figure 10) and a CYP2C9 exon 3 3'-end reverse primer (HF2R, AGCCTTGGTTTTTCTCAACTC, Figure 10) were used at a concentration of 250μM each. Amplifications were performed in a volume of 100 μl containing 20 mM Tris HCl (pH 8.3), 1.5 mM $MgCl_2$, 25 mM KCl, 0.05% (w/v) Tween 20, 10 μg gelatin/ml, 2% (w/v) DMSO, 200 μM each of dATP, dCTP, dGTP and dTTP and 2.5 units of Taq DNA polymerase (Perkin-Elmer). Reactions were carried out for 35 cycles

at an annealing temperature of 55°C for 90 sec, a polymerase temperature of 72°C for 1 min, and a heat denaturing temperature of 93°C for 1 min, using a Perkin-Elmer Cetus DNA thermal cycler. The PCR products digested with *AvaII* and sized using NuSieve agarose gels (3% NuSieve, 0.75% agarose). Presence of the CYP2C9 wild-type and R144C alleles were detected as fragments of 50 + 270 bp and 320 bp respectively (see Figures 3). The PCR product synthesized from human genomic DNA with the primers HF18/HF2R was directly sequenced on an ABI 373A automatic sequencer. Briefly, the PCR product was first purified by using the Wizard DNA clean-up system (Promega Co., Madison, WI). The purified template was then subjected to dideoxy terminator cycle-sequencing with the primers HF18 and HF2R. The primer-extended products were purified and sequenced following the manufacturer's procedure. Sequence analysis was done by using the MacVector software program (Eastman-Kodak Co., Rochester, NY).

DNA was obtained from 94 patients. Of these 58 (62%) were homozygous for the wild-type CYP2C9 gene and 36 (38%) were heterozygous for the R144C allele. No R144C homozygotes were found. The frequency of the wild-type (Arg-144) and R144C (Cys-144) alleles in the study population is thus 0.808 and 0.192 respectively. An expectation of 3.7% R144C homozygotes can be anticipated from the Hardy-Weinberg equilibrium, but the 95% confidence interval in this estimation of 0.8-8.4% and thus the finding of zero homozygotes in 94 patients is not significantly different from expectation. The specificity of the PCR reaction with respect to the CYP2C9 gene was confirmed by sequencing. The alignment of the sequence obtained from the PCR product with that corresponding to the CYP2C9 gene showed a 100% degree of homology. Interestingly, a heterozygous pattern was obtained for the R144C allelic variant, confirming the high frequency of this allele within the normal population. No sequence

deriving from CYP2C9, CYP2C18 or CYP2C19 was found confirming the specificity of the assay for CYP2C9.

EXAMPLE 3

Method for determining the genotype CYP2A6

Genotyping for the CYP2A6 polymorphism is carried out by allele-specific PCR using two parallel PCR reactions, one specific for the wild-type allele, one for the mutant allele. Amplifications are performed in 0.5 ml microcentrifuge tubes in a volume of 45 µl containing 10 mM Tris-HCl, pH 8.8, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 % Triton X-100, 5 % dimethylsulfoxide, 200 µM each of dTTP, dATP, dCTP and dGTP, 250 µM of the primers B and either J51 or J61, 1.25 units Taq polymerase and 1 µg human leukocyte genomic DNA. PCR conditions consist of 40 cycles with a denaturation at 93°C for 1 min., annealing at 57°C for 2 min and polymerization at 70°C for 2 min. The products are analyzed by electrophoresis on 1% agarose minigels in TBE buffer and DNA is visualized by staining with ethidium bromide. As shown in Figure 4, there are three possible results: the individual may be homozygous for the wild-type allele and give a DNA product only for the PCR reaction with primer J51, the individual may be heterozygous with one wild-type and one mutant allele and give DNA products with both primers J51 and J61 or the individual may be homozygous for the mutation and give a DNA product only with the J61 primer.

EXAMPLE 4

Alternative Method for Determining the Genotype CYP2A6

For use of F4 and R4 primers, each reaction mixture contained 600 ng human genomic DNA, 0.2 µM of each primer, 200 µM dNTP's, 0.8 mM magnesium acetate and 2 units of rTth I DNA polymerase. Hot start was as indicated by the manufacturer (Perkin Elmer) and the amplification reaction of 31 cycles of 93°C, 1 min; 66°C, 6 min 30 sec. Amplification products were analyzed in

0.7% agarose gels and the DNA visualized by staining with ethidium bromide. For the exon 3 specific amplification, the reaction which uses, the primers E3F and E3R consist of 5μl of the 7.8 Kb PCR reaction, 0.5 μM of each primer, 200 μM dNTP's, 1.5 mM $MgCl_2$ and 2.5 units of Taq DNA polymerase. The amplification reaction consisted of 94°C for 3 minutes followed by 31 cycles of 94°C, 1 minute; 60°C, 1 minute and 72°C, 1 minute.

Amplification products were then digested without purification with restriction endonucleases which detect the CYP2A6 wild type (no digestion), CYP2A6v1 (*Xcm*I) and CYP2A6v2 (*Dde*I). DNA was visualized by use of ethidium bromide after electrophoresis in 1% agarose, 3% NuSieve agarose.

It is of note that CYP2C9 genotyping can be performed using an allele-specific assay similar to that used above for CYP2A6.

CLAIMS

1. A CYP2A6v2 DNA having a coding sequence shown in Figure 11.

2. The DNA of claim 1 having a genomic sequence as shown in Figure 12.

3. A CYP2A13 DNA having a coding sequence shown in Figure 13.

4. The DNA of claim 3 having a genomic sequence shown in Figure 14.

5. A nucleic acid primer sequence comprising at least ten (10) contiguous nucleotide bases selected from the sequence showing in Figure 12.

6. A nucleic acid primer sequence comprising at least ten (10) contiguous nucleotide bases selected from the sequence shown in Figure 14.

7. A nucleic acid primer sequence selected from the group consisting of:

A. 5′ GGCTTCCTCATCGACGCACT 3′;
B. 5′ GGCTTCCTCATCGACGCACA 3′;
C. 5′ AATTCCAGGAGGCAGGGCCT 3′;
D. 5′ TGCAAGTGCCTGTTTCAGCA 3′;
E. 5′ AGCCTTGGTTTTTCTCAACTC 3′;
F. 5′ CCCCTTATCCTCCCTTGCTGGCTGTGTCCCAAGCTAGGCA GGATTCATGGTGGGGCA 3′;
G. 5′ GCCACCACGCCCCTTCCTTTCCGCCATCCTGCCCCCAGTC TTAGCTGCGCCCCTCTC 3′;
H. 5′ CCTGATCGACTAGGCGTGGTATTCAGCAACGGGGAGCGCG CCAAGCAGCTCCTG 3′;
I. 5′ CGCGCGGGTTCCTCGTCCTGGGTGTTTTCCTTCTCCTGCC CCCGCACTCGGGATGCG 3′;

or any nucleic acid sequence of at least 10 contiguous nucleotides selected from any one of A-I.

8. A method of determining the presence or absence of an allelic variant in CYP2A6 or CYP2C9 DNA comprising:

(a) amplifying an exon containing a variant sequence with in said DNA, producing an extension product;

(b) treating extension products with at least one restriction endonuclease under conditions sufficient to produce digestion fragments;

(c) analyzing the digestion fragments, for a variant specific digestion fragment or lack thereof.

9. The method of claim 8 wherein a CYP2C9 variant DNA is being detected.

10. The method of claim 9 wherein the amplifying step is a polymerase chain reaction using primers comprising HF18 and HF2R.

11. The method of claim 8 wherein step (a) is preceded by a gene-specific amplification reaction.

12. The method of claim 11 wherein the gene-specific amplification is a polymerase chain reaction.

13. The method of claim 12 wherein a CYP2A6 variant is being detected.

14. The method of claim 13 wherein a gene-specific amplification reaction uses primers comprising F4 and R4 and the exon amplification reaction uses primers comprising E3F and E3R.

15. The method according to claim 10 wherein the extension products are treated with the restriction endonuclease AvaII.

16. The method according to claim 14 wherein the extension products are treated with at least one restriction endonuclease comprising *Dde*I and *Xcm*I.

17. A method of determining the presence or absence of an allelic variant in CYP2A6 or CYP2C9 DNA comprising:

(a) contacting said DNA with a first primer encompassing a nucleotide variation specific to variant DNA and a second primer which is complementary to a region of said DNA such that upon hybridization and amplification, an extension product will be formed;

(b) analyzing the extension products for allelic-variant specific extension products.

18. The method of claim 17 wherein a CYP2A6 variant DNA is being detected.

19. The method of claim 18 wherein the amplifying step is a polymerase chain reaction wherein the first primer comprises J51 and J61 and the second primer comprises primer B.

20. A kit for determining the presence or absence of an allelic variant of CYP2A6 or CYP2C9 DNA comprising: at least one nucleic acid primer sequence capable of hybridizing to said DNA; the kit further containing instructions relating to the determination of the presence or absence of an allelic variant of CYP2A6 or CYP2C9 DNA.

21. The kit according to claim 20 further comprising amplification components and at least one restriction endonuclease.

22. The kit of claim 20 wherein the CYP2A6 allelic variant is being detected.

23. The kit of claim 22 wherein the nucleic acid primers comprise F4, R4, E3F and E3R.

24. The kit according to claim 20 wherein the CYP2C9 allelic variant is being detected.

25. The kit according to claim 25 wherein the nucleic acid primers comprise HF18 and HF2R.

26. A process for providing a human with a therapeutic CYP2A6v2 or CYP2A13 DNA segment said human cells expressing in vivo in said human or therapeutically effective amount of said protein.

27. A pharmaceutical composition comprising an antisense nucleic acid derived from CYP2A6v2 DNA.

28. A pharmaceutical composition comprising and antisense nucleic acid derived from CYP2A13.

	..................exon 2..
IIC1 (C9)	GATCTTGGAGAGGAGTTTTCTGGAAGAGGCATTTTCCCACTGGCT
	Asp Leu Gly Glu Glu Phe Ser Gly Arg Gly Ile Phe Pro Leu Ala
IIC2 (C8)	GATAATGGAGAGGAGTTTTCTGGAAGAGGCAATTCCCCAATATCT
	Asp Asn Gly Glu Glu Phe Ser Gly Arg Gly Asn Ser Pro Ile Ser
DIIC2	
Clone 4 (hIIc1-4)	GATCTTGGAGAGGAGTTTTCTGGAAGAGGCCATTTCCCACTGGCTG
Clone 18 (hIIC1-18)	GATCATGGAGAGGAGTTTTCTGGAAGAGGAAGTTTTCCAGTGGCTG
Clone 3	
Clone 16	GATCATGGAGAGGAGTTTTCTGGAAAAGGTATTTTCCCAGTATCCA
Clone 21 (hIIC1-21)	GATCTTGGAGAGGAGTTTTCTGGAAGAGGCCATTTCCCACTGGCTG
Clone 26 (hIIc1-26)	GATCATGGAGAGGAGTTTTCTGGAAGAGGAAGTTTTCCAGTGGCTG
Clone 33	

FIG. 1 (Sheet 1)

```
             ............exon 2.......................][Start of intron 2...........................................
IIC1         GAAAGAGCTAACAGAGGATTTG
             Glu  Arg Ala Asn  Arg Gly Phe G
IIC2         CAAAGAATTACTAAAGGACTTG
             Gln  Arg  Ile Thr   Lys Gly Leu G

DIIC2                                CTTGGTAGGTGCACATATTTCTGTGTCAGCTTTGGTAAC

Clone 4      AAAGAGCTAACAGAGGATTTGGTAGGTGTGCAAGTGCCTGTTTCAGCATCTGTCTTGG
                                          Primer HF-18
Clone 18     AAAAAGTTAACAAAGGACTTGGTAAATGTGCATGTATCGTGTGTATGTGTACATGT

Clone 16     AAAAAGCTA  GTAAGGAGTTGGTACATGTGTGTCAGTGTGTGTGTGCCTTTGTCTG
Clone 21     AAAGAGCTAACAGAGGATTTGGTAGGTGTGCAAGTGCCTGTTTCAGCATCTGTCTTGG
Clone 26     AAAAAGTTAACAAAGGACTTGGTAAATGTGCATGTATCGTGTGTATGTGTACATGT
```

FIG. 1 (Sheet 2)

................Intron 2..

```
DIIC2      TGGGGTGAGGGGGATGGAAAACAGAGCCCTAAAAAGCTTCTCAGCAGAGCTTAGC
Clone 4    GGATGGGGAGGATGGAAAACAGAGACTTACAGAGCTCCTCGGGCAGAGCTTGGCCCA
Clone 18   GTATGTACTGGGCAGTGGCTATAGGGATGGGGAGGATGGAAAACAGGCTTGAAAA
Clone 3                                  CAGAAGGTGAAT(G)GAAACAACAC(T)TGAA
Clone 16   TATTAGTAATGAGGCAGAAGGTGAATGGAAAACAAACACTTGAAGAGCTCCTAAA
Clone 21   GGATGGGGAGGATGGAAAACAGA[               CTA   GCAGAGCT(T)]CTCGGG
Clone 26   GTATGTACTGGGCAGTGGCTATAGGGATGGGGAGGATGGAAAACAGGCTTGAAGA
```

FIG. 1 (Sheet 3)

......................Intron 2..

DIIC2	CTATCTGCATGGCTGCCAAGTGTTGCAGCACTTTCTTCCTTGGCTGTGAATTCTC
Clone 4	TCCACATGGCTGCCCAGTGTCAGCTTCCTCTTTCTTGCCTGGGATCTCCCTCCTA
Clone 18	GCTCCTGGGACAGAACTTGACCTGTCCACGTGGCTGCCGAGTGTCAGCTCTCTTG
Clone 3	GAGCTCCTAAAC(T)TAGC(T)TAGCTTGGCCATTGGGTGGCTGTTGAAAATCAGCTTC
Clone 16	ACTTAGCTTGGCC(C)ATTTGGTGGCTGTTGAAATCAGCTTCCTCTTTCNNNC(C)TGG
Clone 21	CAGAGCTTGGCCCATCCACATGGCTGCCCAGTGTCAGCTTCCTCTTTCTTGCCTG
]Clone 26	GCTCCTGGGACAGAACTTGACCTGTCCACGTGGCTGCCGAGTGTCAGCTCTCTTG

FIG. 1 (Sheet 4)

....end of intron 2]

DIIC2	CCAGTTTCTGCCCCTTTTTTTATTAG	
Clone 4	GTTTCGTTTCTCTTCCTGTTAG	
Clone 18	TCCTTGTTTGGATTCTCCCTCGTAGCTTCTGTTTTCTGTTCTGCTAG	
Clone 3	CTCTTTCTTGCCTGGGATCTCCCTCCTCGTTTCTGTTTCCCTTCCTTTCA	
Clone 16	ATCTCCTCCTCGTTTCTGTTCCTCCTTC	A
Clone 21	GGATCTCCCTCCTAGTTTCGTTTCTCTTCCTGTT	AG
Clone 26	TCCTTGTTTGGATTCTCCCTCGTAGCTTCTGTTTTCTGTTCTGCTAG	

FIG. 1 (Sheet 5)

```
            [Start of exon 3..........................................................

IIC1        GAATTGTTTTCAGCAATGGAAAGAAATGGAAGGAGATCCGGCGTTTCTCCCTCATGACG
            ly Ile  Val Phe Ser Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu Met Thr
IIC2        GAATCATTTCCAGCAATGGAAAGAGATGGAAGGAGATCCGGCGTTTCTCCCTCACAACC
            ly Ile  Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu Thr Thr

DIIC2       GAATCATTTCCAGCAATGGAAAGAGATGGAAGGAGATCCGGCGTTTCTCCCTCACAACC

Clone 4     GAATTGTTTTCAGCAATGGAAAGAAATGGAAGGAGATCAGGCGTTTCTCCCTCATGACG

Clone 18    GAATCCTTTTCAGCAATGGAAAGAGATGGAAGGAGATCCGGCGTTTCTGCCTCATGACT

Clone 3     GGATCATTTTTAGCAATGGAAAGAGATGTAAGGATGTCTGGCTCTTCTTGCTCATGACG

Clone 16    GGATCATTT

Clone 21    GAATCGTTTTCAGCAATGGAAAGAGATGGAAGGAGATCCGGCGTTTCTCCCTCATGACG

Clone 26    GAATCCTTTTCAGCAATGGAAAGAGATGGAAGGAGATCCGGCGTTTCTCCCCCATGACG
Clone 33                                                    G  T      T
```

FIG. 1 (Sheet 6)

```
             ............exon 3.....................................................................
IIC1         CTGCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGACTGTGTTCAAGAGGAAGCCCG
             Leu Arg Asn Phe Gly Met Gly Lys  Arg Ser Ile  Glu  Asp Cys Val  Gln Glu  Glu  Ala Ar
IIC2         TTGCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGACCGTGTTCAAGAGGAAGCTCA
             Leu Arg Asn Phe Gly Met Gly Lys  Arg Ser Ile  Glu  Asp Arg Val  Gln Glu  Glu  Ala Hi
                                                          ↑
                                          Site of A144C polymorphism

DIIC2        TTGC

Clone 4      CTGCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGACCGTGTTCAAGAGGAAGCCCG

Clone 18     CTGCGGAATTTTGGGATGGGGAAGAGGAGCATCGAGGACCGTGTTCAAGAGGAAGCCCG

Clone 3      CTCTGGAATTGTAGGATGGTGAAGAGGAGCAATGGAGA     TGTTCAAGGTGAAGCCCA
                                                  AGCA

Clone 21     CTGCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGACCGTGTTCAAGAGGAAGCCCG

Clone 26     CTGCGGAATTTTGGGATGGGGAAGAGGAGCATTGAGGACCGTGTTCAAGAGGAAGCCCG
Clone 33                                     C
```

FIG. 1 (Sheet 7)

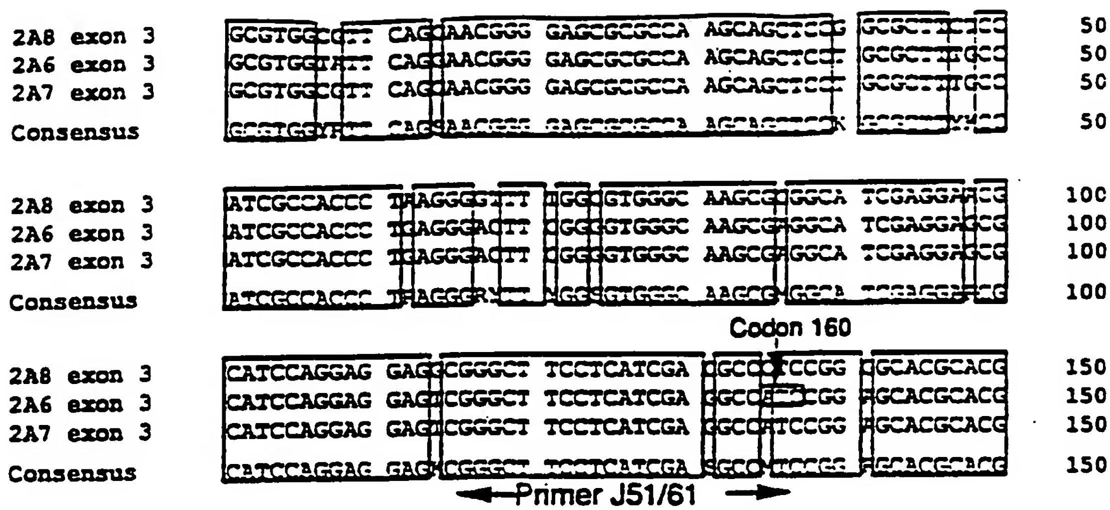
```
2A6 intron 2   --------GT GCCCAAGAGG GGGAAGGTGG GCAGGCGGAC ACGAAGGTCC   42
2A8 intron 2   GTGAGGGGGT G-CCAAGAGG GGGAAGGTGG GCAGGTGGAT CCGAAGGTCC   49
2A7 intron 2   --------GT G-CCAAGAGG GGGAAGGTGG GCAGGCGGAC ACGAAGGTCC   41

Consensus      GTGAGGGGGT GCCCAAGAGG GGGAAGGTGG GCAGGYGGAY MCGAAGGTCC   50

2A6 intron 2   CAGTGTTGCC AGCCTTCTCC CTGACTCTCC TGACAACTGG AGGTTAGGGG   92
2A8 intron 2   CGGTGTCGCC AGCCTTCTCC CTGACTCTCC TGCCCACTGG AGGTCAGGGG   99
2A7 intron 2   CAGTGTTGCC AGCCTTCTCC CTGACTCTCC TGACCACTGG AGGTTAGGGA   91

Consensus      CRGTGTYGCC AGCCTTCTCC CTGACTCTCC TGMCMACTGG AGGTYAGGGR  100

2A6 intron 2   AGAGTCCCCA GCCTGGTCTT CCCTCCCCAT CTCCCTGCAT TGGGGCCTCC  142
2A8 intron 2   AGAGGCCCCT GGCTGGTCTT CTCTCCCCAT CTCCCTGCAT CGGGGTCTCC  149
2A7 intron 2   AGAACCCCCG GCCTGGTCTT ATCTCCCCAT CTCCCTGCAC TGGGGCCTCC  141

Consensus      AGARYCCCCD GSCTGGTCTT HYCTCCCCAT CTCCCTGCAY YGGGGYCTCC  150

2A6 intron 2   CCATGTGTAT CCCGCACCTG TCTCCAGCCG CCCACTCCTG ATTCCTCCCT  192
2A8 intron 2   CCGTGTGAGT CCCACACCTG TCTCCAGCCC CCCTCCCCTG ATTCCTCCCT  199
2A7 intron 2   CCATGTGTAT CCCCCACCTG TCTCCAGCCC CCCGCTCCTG ATTCCTCCC-  190

Consensus      CCRTGTGWRT CCCNCACCTG TCTCCAGCCS CCCNCYCCTG ATTCCTCCCT  200

2A6 intron 2   GGCTCTCTCT GCCCCAGCTC CTTATTCTCT CTCAGTGGAG TCTCCTCTTT  242
2A8 intron 2   GGCTCTCTCT GCCCCGTCTC CTCCTTCTCT CTCAGTGGAG TCTCCTCTTA  249
2A7 intron 2   GGCTCTCTCT GCCCCAGCTC CAGATTCTCT CTCA-TGGAG TCTCCTCTTT  239

Consensus      GGCTCTCTCT GCCCCRKCTC CWBMTTCTCT CTCAGTGGAG TCTCCTCTTW  250

2A6 intron 2   GGCCTCTCTC TCCAGCTCCA AGGACATCCT GGGTTTCTGT TTACCAGCCC  292
2A8 intron 2   GGCCTCTCTC TCCAGCTCTG AGGACATCCG GGGTTTCTGT TTACCAGCCC  299
2A7 intron 2   GTCCTCTCTC TCCAACTCTG AGGACATC-- ---------- ----------  267

Consensus      GKCCTCTCTC TCCARCTCYR AGGACATCCX GGGTTTCTGT TTACCAGCCC  300

2A6 intron 2   TGGGTCTCTG TCTACATGAG TCTTTGACGC CCTCTTAGCT TCTGGGCTTC  342
2A8 intron 2   TGGTCCTCTG TCTTCATTTG TCTTTTTGTC GCTCTCGGCT TCTGTGCTTC  349
2A7 intron 2   ---------- ---------- -----TTCG- ---------- ----------  271

Consensus      TGGKYCTCTG TCTWCATKWG TCTTTKAGKC SCTCTYRGCT TCTGKGCTTC  350

2A6 intron 2   TCTGGGTTTC TCATCTCTCC GGATCCCTTT CTCAATTCTT CCTCTGTCTT  392
2A8 intron 2   TCCGTGTTTC TCCTCTCTCT GCTTCCCTCT CCCACTTCTT CCTCTGTCTT  399
2A7 intron 2   ---------- ---------- ---------- ---------- ----------  271

Consensus      TCYGKGTTTC TCMTCTCTCY GSWTCCCTYT CYCAMTTCTT CCTCTGTCTT  400

2A6 intron 2   AGGATGCCAG GGTTATTCCT ACTTCCACAT CTTCAGGCTC CATCTCCTGG  442
2A8 intron 2   AGGATTTCAG GGT-ATTCCT ACTTCCACAT CTCCAGCTCC CAACTCCTGG  448
2A7 intron 2   ---------- ---------- ---------- ---------- ----------  271

Consensus      AGGATKYCAG GGTTATTCCT ACTTCCACAT CTYCAGSYYC CAWCTCCTGG  450

2A6 intron 2   TAACAGTCTC TCTTCCTTCC AGACCCTCTC TGTTTCTATC TCAATATTAA  492
2A8 intron 2   TAATTGTCTG TCCTCCTTCC CGATCCTCTC TGTTTCTGTC TCCATATTTT  498
2A7 intron 2   ---------- ---------- ---------- ---------- ----------  271

Consensus      TAAYWGTCTS TCYTCCTTCC MGAYCCTCTC TGTTTCTRTC TCMATATTWW  500

2A6 intron 2   ACTCTCT--G CTCCAGCTCA GCTTAAGAAT CTCACACCAA GAGAGGATGT  540
2A8 intron 2   TCTCTCTCTT CTCCAGTTCA GATTAAGAAT CTTTCACCA- -TTTTTATTT  546
2A7 intron 2   ---------- ---------- ---------- ---------- ----------  271

Consensus      WCTCTCTCTK CTCCAGYTCA GMTTAAGAAT CTYWCACCAA GWKWKKATKT  550
```

FIG. 2 (Sheet 1)

```
2A6 intron 2   CCTCCACCCA GATCTCCCCA TATCTCACTA CCCCACCCTC CATC---CTC   587
2A8 intron 2   CCTCCTCCCA GATCTCCCCA TATCTCACTT CCCCTCCCTC CATCTCTCTC   596
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      CCTCCWCCCA GATCTCCCCA TATCTCACTW CCCCWCCCTC CATCTCTCTC   600

2A6 intron 2   TGCCT----C CATCAC--TC TCTTTCTC-- -------TCC CC--A-----   615
2A8 intron 2   TTTCTCTCCC CACTACCTTC CCTTCCTCCA TGGAGTATCC CCGTATCCCT   646
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      TKYCTCTCCC CAYYACCTTC YCTTYCTCCA TGGAGTATCC CCGTATCCCT   650

2A6 intron 2   CTGCCCCTGC GGACGCGATC CAATGG--AG TGTG------ ----GA---G   650
2A8 intron 2   CTGTTTCTCT GCATCTGTCT GTCTGGCCTT TCTGCTTCTC TTCTGATTCT   696
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      CTGYYYCTSY GSAYSYGWYY SWMTGGCCWK TSTGCTTCTC TTCTGATTCK   700

2A6 intron 2   CTAATGCCGT -------GAA GCTATGTGCA TCTCTCTGTC TGGCCGTACC   693
2A8 intron 2   CTTATTCTTT CTACCCGGAC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC   746
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      CTWATKCYKT CTACCCGGAM KCTMTSTSYM TCTCTCTSTC TSKCYSTMYC   750

2A6 intron 2   TGGGT---AA TAACCTGATC GACT------ ---------- ----------   714
2A8 intron 2   TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTA TATATATATA   796
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      TSKSTCTCWM TMWCYYKMTC KMYYTCTCTC TCTCTCTCTA TATATATATA   800

2A6 intron 2   ---------- ---------- ---------- ---------- ----------   714
2A8 intron 2   TATATATATA CACACACACA CACACACACA CACACACACA CACACACATA   846
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      TATATATATA CACACACACA CACACACACA CACACACACA CACACACATA   850

2A6 intron 2   ---------- ---------- ---------- ---------- ----------   714
2A8 intron 2   TATATTAGGG GGGGACTCCC TTTCTGCTCC ACCCTTGGGG AGCCCCTTGG   896
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      TATATTAGGG GGGGACTCCC TTTCTGCTCC ACCCTTGGGG AGCCCCTTGG   900

2A6 intron 2   ---------- ---------- ---------- ---------- ----------   714
2A8 intron 2   AACTGGTCCG CTCTGCTACC ACCACCCCCT GACCTCTCTC CACCCCCGCG   946
2A7 intron 2   ---------- ---------- ---------- ---------- ----------   271

Consensus      AACTGGTCCG CTCTGCTACC ACCACCCCCT GACCTCTCTC CACCCCCGCG   950

2A6 intron 2   ---------- --                                            714
2A8 intron 2   TTCACCTCCC CA                                            958
2A7 intron 2   ---------- --                                            271

Consensus      TTCACCTCCC CA                                            962
```

Intron 2 alignment

FIG. 2 (Sheet 2)



Exon 3 alignment

FIG. 2 (Sheet 3)

```
2A8 intron 3   ---------- ---------- ---------G TGAGTAGGGG ACGCCGAGTG    21
2A6 intron 3   GTGAGCAGGG GACCCCGAGT GCGGGGGCAG AGAGAAGGAA AACAC-----    44
2A7 intron 3   GTGAGTAAGG TTCCCCGAGT GCGGGGGCAG AGAGAAGGAA AACAC-----    44
Consensus      GTGAGYARGG KWCCCCGAGT GCGGGGGCAG NGAGNAGGRR RNCNCGAGTG    50

2A8 intron 3   CGAGGGCG-G GAACC--CGC GCTTTCTGCC TGGGGATGGG GACTAGGTGG    68
2A6 intron 3   CCAGGACGAG GAACCCGCGC GCCTTCTGCC TGCGGATGGG GACTAGGTGG    94
2A7 intron 3   CCAGGACGAG GAACCCGCGC GCCTTCTGCC TGCGGATGGG GACTAGGTGG    94
Consensus      CSAGGRCGRG GAACCCGCGC GCYTTCTGCC TGSGGATGGG GACTAGGTGG   100

2A8 intron 3   GGAAAGGGGC CCGCACTTCC AGCCCTGGAG TCTGGCGCT- GGCATTCGGC   117
2A6 intron 3   GGAAAGGCGC CCGCACTTCC AGCCCTGGAG TCTGGCGCTG GGAATTTGGC   144
2A7 intron 3   GGAAAGGCGC CCGCACTTCC AACCCTGGAA TCTGGCGCTG GGCATTTGGC   144
Consensus      GGAAAGGSGC CCGCACTTCC ARCCCTGGAR TCTGGCGCTG GGMATTYGGC   150

2A8 intron 3   TCAACAGGGC CCTGCCTCCT GGAATTCTGA CTCTCCTCAG ACCTCTGAGT   167
2A6 intron 3   TCAACAAGGC CCTGCCTCCT GGAATTCTGA CTCTCCTCAG ACCTCTGAGT   194
2A7 intron 3   TCAACAAGGC CCTGCCTCCT GGAATTCTGA CTCTCCTCAG ACCTCTGAGT   194
Consensus      TCAACARGGC CCTGCCTCCT GGAATTCTGA CTCTCCTCAG ACCTCTGAGT   200
                   <---------Primer B--------->

2A8 intron 3   TGACTCTCTC CCCAACCCCC CTTCTCCCGC CACACCTGTA -   207
2A6 intron 3   TGACTCTCTC CCCAACCCCC T-TCTCCCGA CATACCCGCA -   233
2A7 intron 3   TGACTCTCTC CCCAACCCCC TCCCTCCCTC CACACCTGCA G   235
Consensus      TGACTCTCTC CCCAACCCCC YYYCTCCCKM CAYACCYGHA G   241
```

Intron 3 alignment

FIG. 2 (Sheet 4)

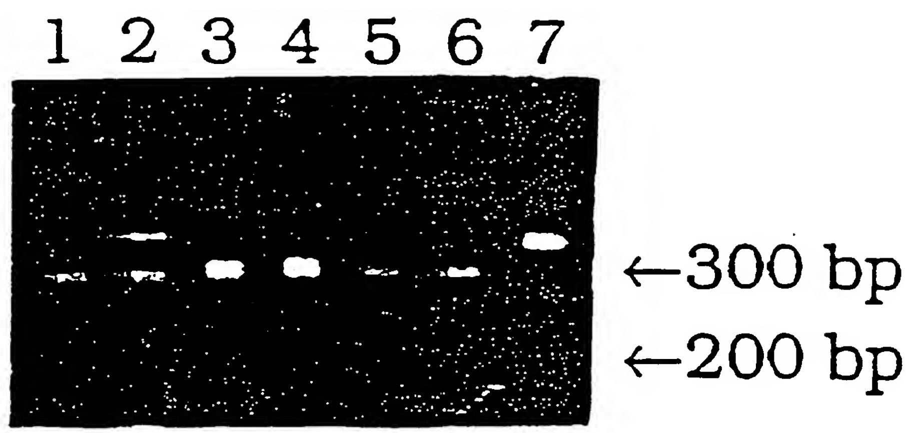


FIG. 3

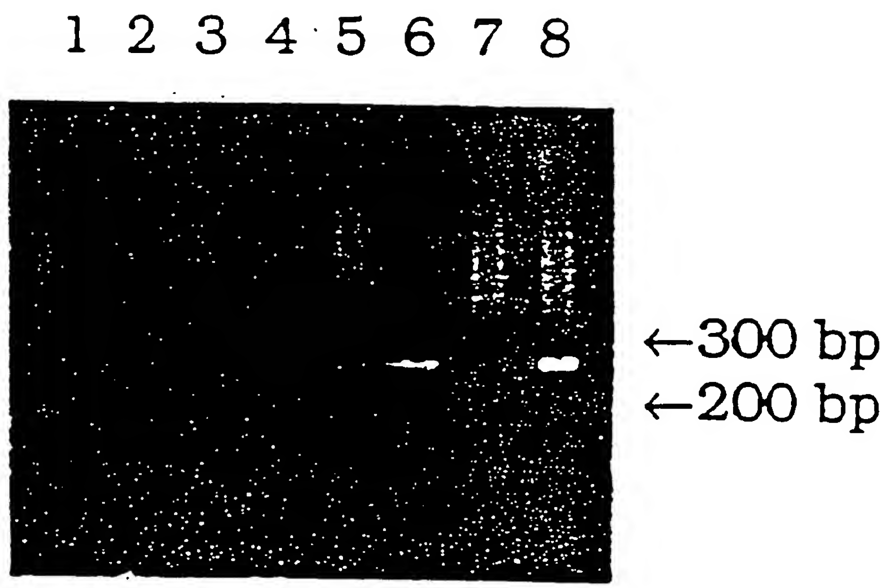
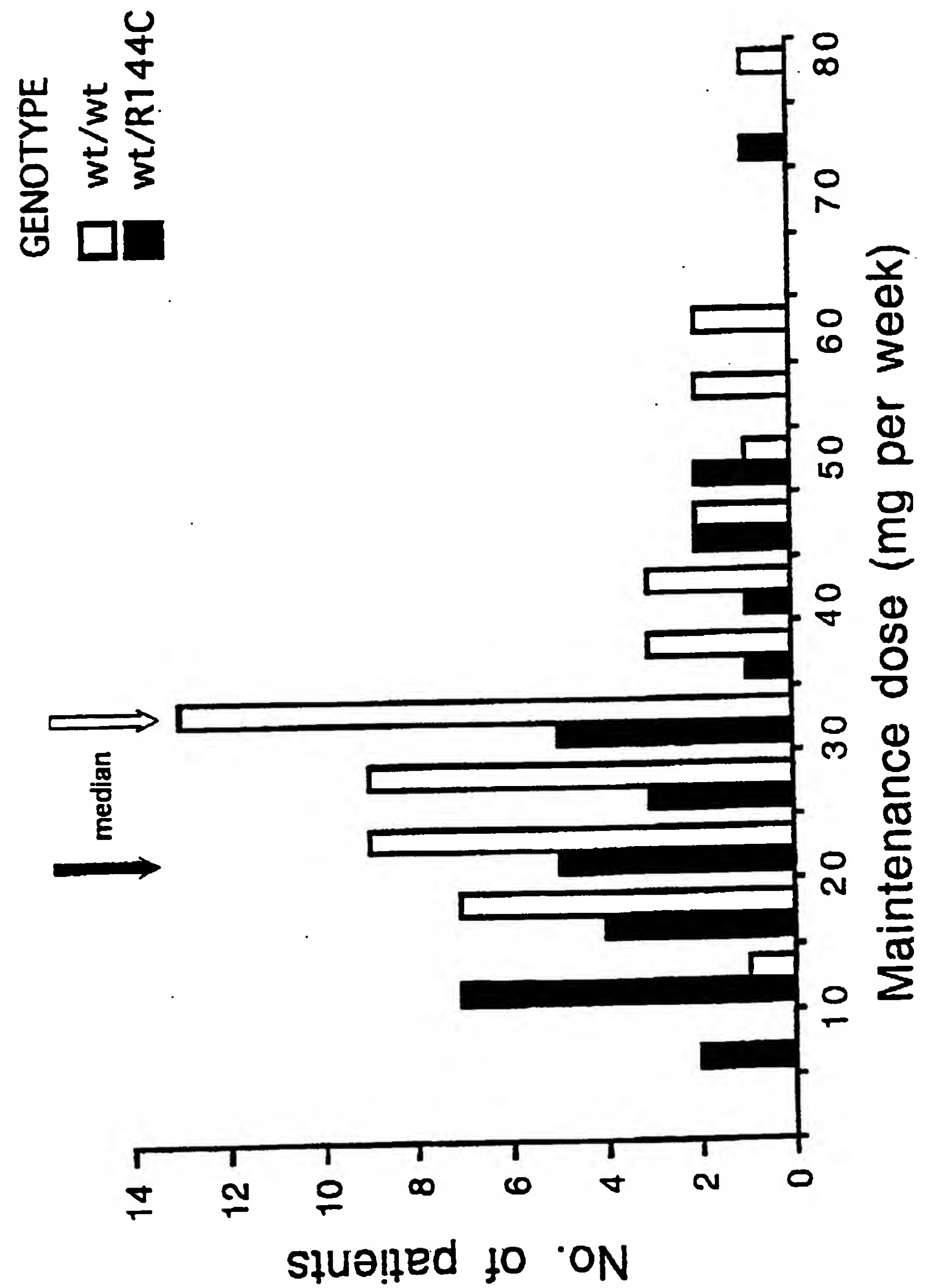
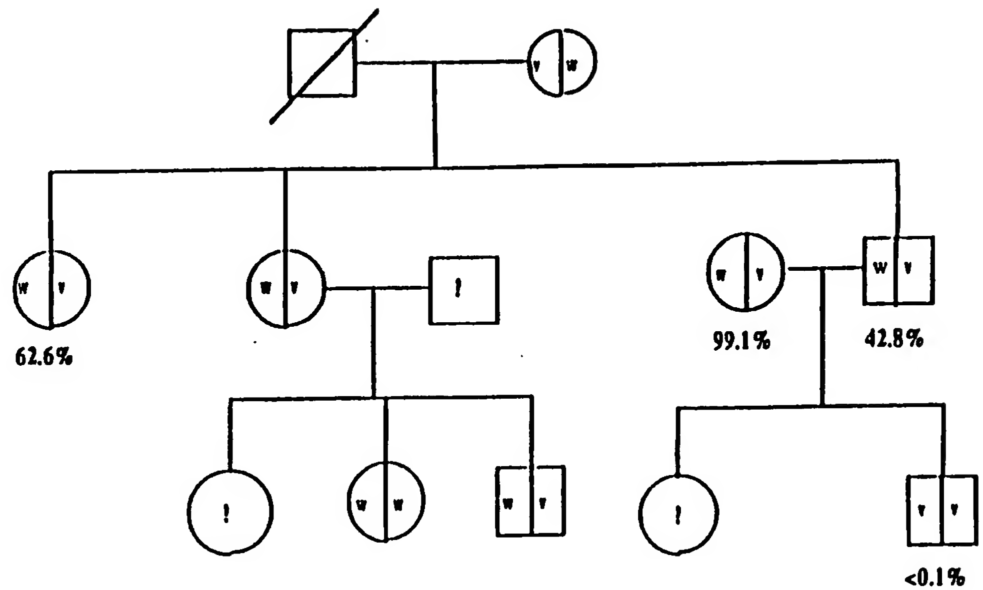


FIG. 4

FIG. 5


GENOTYPE
wt/wt
wt/R144C
median
No. of patients
0
2
4
6
8
10
12
14
Maintenance dose (mg per week)
10
20
30
40
50
60
70
80

7-Hydroxylation of coumarin (%) in a family genotyped for the presence of *CYP2A6* and *CYP2A6v* alleles, showing subject homozygous for *CYP2A6v* who is deficient in coumarin 7-hydroxylation



w = *CYP2A6* wild-type
v = *CYP2A6v* mutant allele
? = not determined

FIG. 6

```
2A6 cDNA  GCGTGGTATTCAGCAACGGGGAGCGCGCCAAGCAGCTCCGGCGCTTCTCCAT
2A6 gene                                         T      TG
2A6 cDNA  CGCCACCCTGCGGGACTTCGGGGTGGGCAAGCGAGGCATCGAGGAGCGCATC
2A6 gene            A              C
2A6 cDNA  CAGGAGGAGGCGGGCTTCCTCATCGACGCCCTCCGGGGCACTGGC
2A6 gene            T                G   A     A   GCA
```

Comparison of *CYP2A6* cDNA and genomic sequences for exon 3

FIG. 7

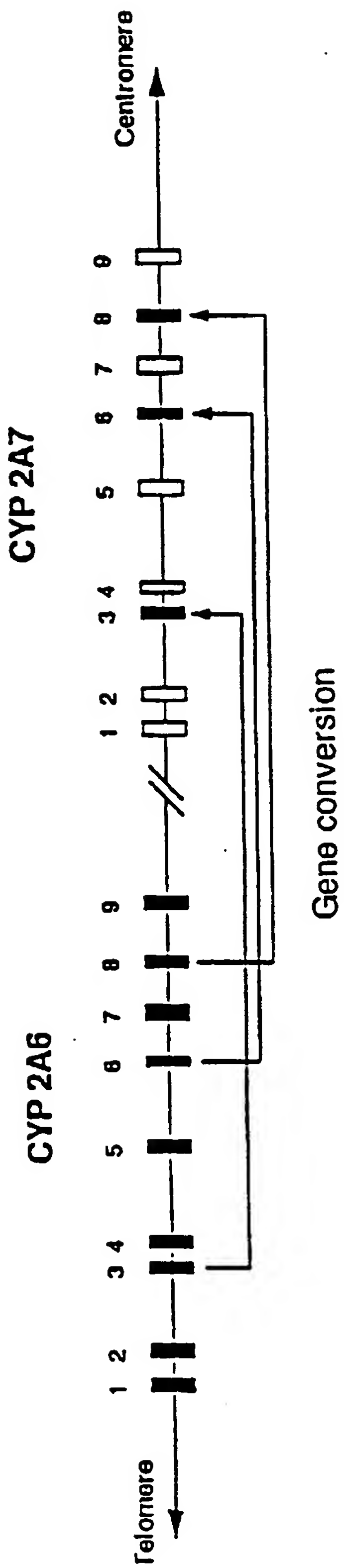

Centromere
CYP 2A7
1 2 3 4 5 6 7 8 9
CYP 2A6
1 2 3 4 5 6 7 8 9
Gene conversion
Telomere

FIG. 8A

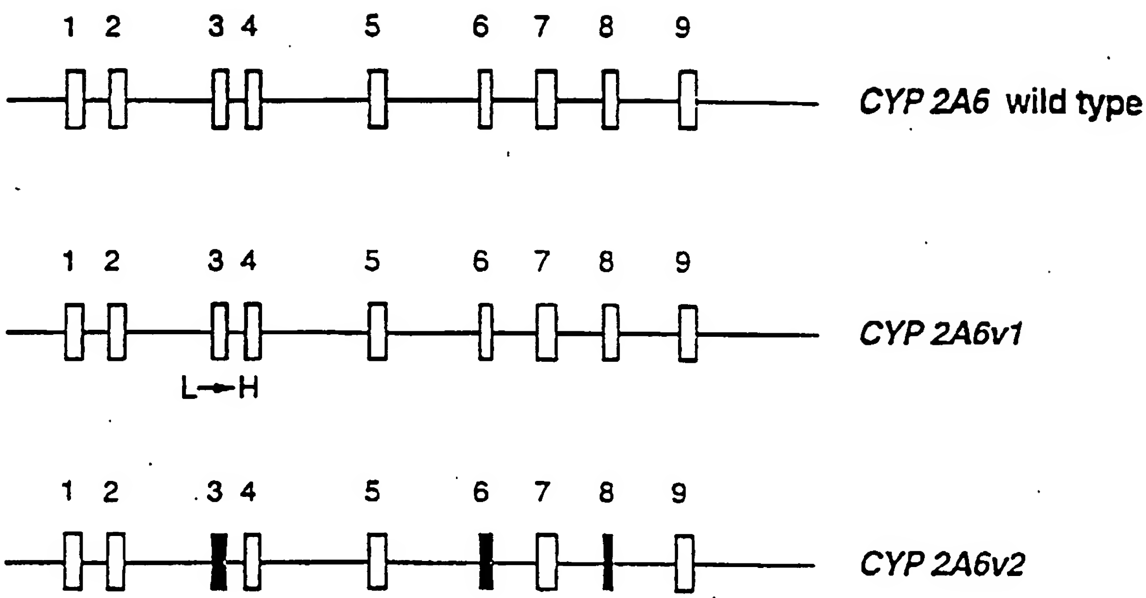

1 2 3 4 5 6 7 8 9
CYP 2A6 wild type
1 2 3 4 5 6 7 8 9
CYP 2A6v1
L→H
1 2 3 4 5 6 7 8 9
CYP 2A6v2

FIG. 8B

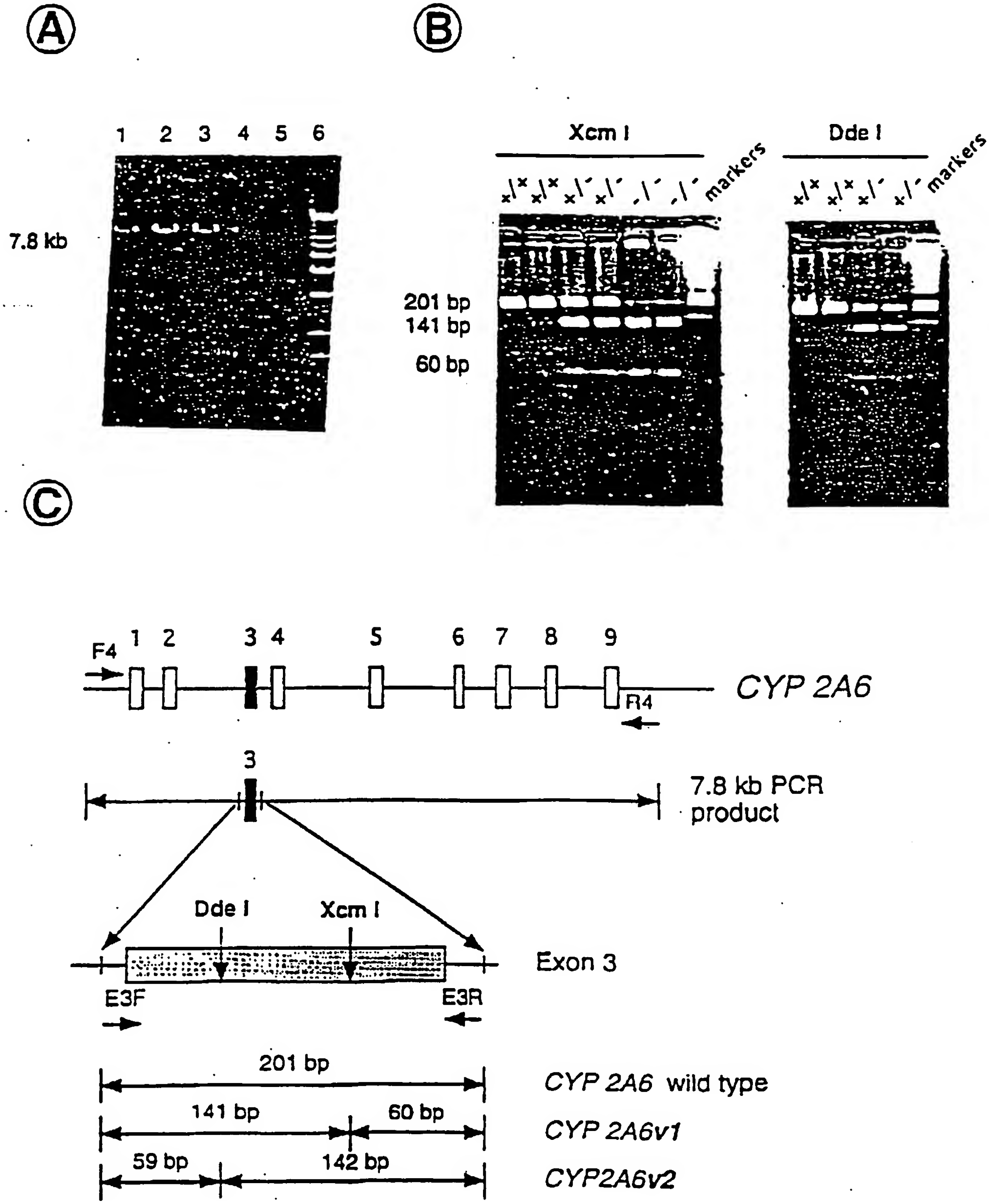

A
B
C
1 2 3 4 5 6
7.8 kb
Xcm I
Dde I
markers
201 bp
141 bp
60 bp
F4
R4
CYP 2A6
7.8 kb PCR product
Dde I
Xcm I
Exon 3
E3F
E3R
201 bp
141 bp
60 bp
59 bp
142 bp
CYP 2A6 wild type
CYP 2A6v1
CYP2A6v2

FIG. 9

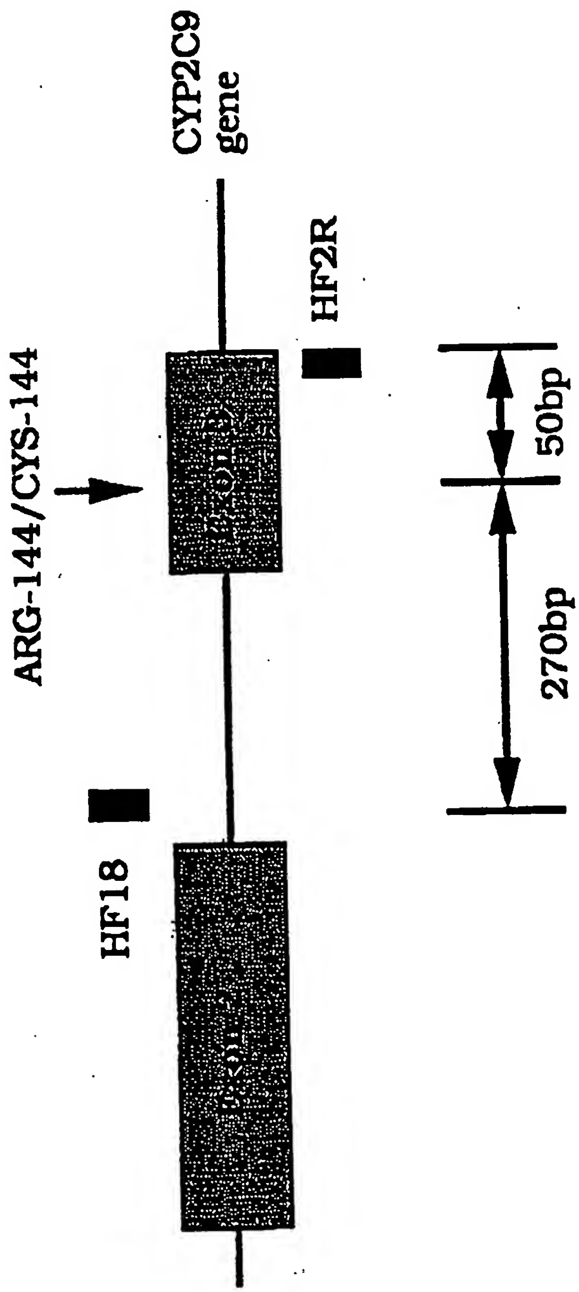


FIG. 10

CYP2A6v2 cDNA.

5'TCTACCACCATGCTGGCCTCAGGGATGCTTCTGGTGGCCTTGCTGGCCTGCCT
GACTGTGATGGTCTTGATGTCTGTTTGGCAGCAGAGGAAGAGCAAGGGGAA
GCTGCCTCCGGGACCCACCCCATTGCCCTTCATTGGAAACTACCTGCAGCTGA
ACACAGAGCAGATGTACAACTCCCTCATGAAGATCAGTGAGCGCTATGGCC
CCGTGTTCACCATTCACTTGGGGCCCCGGCGGGTCGTGGTGCTGTGTGGACATG
ATGCCGTCAGGGAGGCTCTGGTGGACCAGGCTGAGGAGTTCAGCGGGCGAGGC
GAGCAAGCCACCTTCGACTGGGTCTTCAAAGGCTATGGCGTGGTATTCAGCA
ACGGGGAGCGCGCCAAGCAGCTCCTGCGCTTTGCCATCGCCACCCTGAGGGACT
TCGGGGTGGGCAAGCGAGGCATCGAGGAGCGCATCCAGGAGGAGTCGGGCTTC
CTCATCGAGGCCATCCGGAGCACGCACGGCGCCAATATCGATCCCACCTTCTTC
CTGAGCCGCACAGTCTCCAATGTCATCAGCTCCATTGTCTTTGGGGACCGCTT
TGACTATAAGGACAAAGAGTTCCTGTCACTGTTGCGCATGATGCTAGGAAT
CTTCCAGTTCACGTCAACCTCCACGGGGCAGCTCTATGAGATGTTCTCTTCGG
TGATGAAACACCTGCCAGGACCACAGCAACAGGCCTTTCAGTTGCTGCAAGG
GCTGGAGGACTTCATAGCCAAGAAGGTGGAGCACAACCAGCGCACGCTGGA
TCCCAATTCCCCACGGGACTTCATTGACTCCTTTCTCATCCGCATGCAGGAGG
AGGAGAAGAACCCCAACACGGAGTTCTACTTGAAGAACCTGATGATGAGC
ACGTTGAACCTCTTCATTGCAGGCACCGAGACGGTCAGCACCACCCTGCACTA
TGGCTTCTTGCTGCTCATGAAGCACCCAGAGGTGGAGGCCAAGGTCCATGAG
GAGATTGACAGAGTGATCGGCAAGAACCGGCAGCCCAAGTTTGAGGACCGG
GCCAAGATGCCCTACATGGAGGCAGTGATCCACGAGATCCAAAGATTTGGA
GACGTGATCCCCATGAGTTTGGCCCGCAGAGTCAAAAAGGACACCAAGTTTC
GGGATTTCTTCCTCCCTAAGGGCATAGAAGTGTTCCCTATGTTGGGCTCCGTG
CTGAGAGACCTCAGGTTCTTCTCCAACCCCCGGGACTTCAATCCCCAGCACTTC
CTGGGTGAGAAGGGGCAGTTTAAGAAGCGTGATGCTTTTGTGCCCTTCTCCA
TCAGAAAGCGGAACTGTTTCGGAGAAGGCCTGGCCAGAATGGAGCTCTTTCT
CTTCTTCACCACCGTCATGCAGAACTTCCGCCTCAAGTCCTCCCAGTCACCTA
AGGACATTGACGTGTCCCCCAAACACGTGGGCTTTGCCACGATCCCACGAAA
CTACACCATGAGCTTCCTGCCCCGCTGAGCGAGGGCTGTGCCGGTGAAGGTCTG
GTGGGCGGGGCCAGGGAAAGGGCAGGGCCAAGACCGGGCTTGGGAGAGGGGC
GCAGCTAAGACTGGGGGCAGGATGGCGGAAAGGAAGGGGCGTGGTGGCTAG
AGGGAAGAGAAGAAACAGAAGCGGCTCAGTTCACCTTGATAAGGTGCTTCC
GAGCTGGGATGAGAGGAAGGAAACCCTTACATTATGCTATGAAGAGTAGT
AATAATAGCAGCTCTTATTTCCTGA 3'

FIG. 11

```
   1 AAGTTCCCCT GAAATATGGC TCTGGTCTTC CTCCCCTTGC CAATGAAGAA GATGGCAGTG
  61 GAGGTTCTAT GGCAGCCATC CTGGCCTCAC TCTGAGGTTC CAATGAGGAT TCTGGGCATC
 121 AAGAGACAGC TCTGGGCAAA GCTAAATCAA GTCAGCCCCT GGACCCAGTG CTGGGCTGCT
 181 GGGCTTTCTG GGAGAACGCC GCTGGGCTTG CTACACACTC CTCCTCCCAG AAACTCCACA
 241 CCCACAGCCC TGGGTCTTCC TAGCCCCGAG ACTTTCAAGT CCATATGCCT GGAATCCCCC
 301 TTCCTGAGAC CCTTAACCCT GCATCCTCCA CAACAGAAGA CCCCTAAATG CACAGCCACA
 361 CTTTGTCTTA CCCTAATAAA ACCCAGACCT TTGGATTCCT CTCCCCTGGA ACCCCCAGAT
 421 CCGCACAACT TTGGGGTGCA TTCTCACTCT CAGACCCCAA ATCCAAAGCC CAAGTGCTCC
 481 CCTATGCAAA TATTCCAAAC TCCTCAGTTC TACAGCTTAT CTGTTGCCCC CTCCTAAATC
 541 CACAGCCCTG CGGCACCCCT CCTGAAGTAC CACAGATTTA GTCTGGAGGC CCCCTCTCTG
 601 TTCAGCTGCC CTGGGGTCCC CTTATCCTCC CTTGCTGGCT GTGTCCCAAG CTAGGCAGGA
 661 TTCATGGTGG GGCATGTAGT TGGGAGGTGA AATGAGGTAA TTATGTAATC AGCCAAAGTC
 721 CATCCCTCTT TTTCAGGCAG TATAAAGGCA AACCACCCCA GCCGTCACCA TCTATCATCC
 781 CTCTACCACC ATGCTGGCCT CAGGGATGCT TCTGGTGGCC TTGCTGGCCT GCCTGACTGT
 841 GATGGTCTTG ATGTCTGTTT GGCAGCAGAG GAAGAGCAAG GGGAAGCTGC CTCCGGGACC
 901 CACCCCATTG CCCTTCATTG GAAACTACCT GCAGCTGAAC ACAGAGCAGA TGTACAACTC
 961 CCTCATGAAG GTGTCCCAAG ACAGGGAGAT GGGTGTCTCG GGGTGGGGGC TGCCTAGTTG
1021 GCTGGGGCTT TGTGGCAGGG GGTTGACCAG TGTGGACCAG AGTCTTAGGA AATGGAGTTT
1081 TGGAGTTTCA GCATCAGAAA GACAGGATCT TGGGATGTCC AGCTCCCTGA CTGTGAGAAC
1141 CTGGGTGCGA AGCATCCCAG CACATGACAT CTCGGTGCTG GGCCCCATTC AGAGTGGAGG
1201 GTTCTCCCTC TAACCACTCC CACCCACCTC CATCAGATCA GTGAGCGCTA TGGCCCCGTG
1261 TTCACCATTC ACTTGGGGCC CCGGCGGGTC GTGGTGCTGT GTGGACATGA TGCCGTCAGG
1321 GAGGCTCTGG TGGACCAGGC TGAGGAGTTC AGCGGGCGAG GCGAGCAAGC CACCTTCGAC
1381 TGGGTCTTCA AAGGCTATGG TGCCCAAGAG GGGGAAGGTG GGCAGGTGGA CACGAAGGTC
1441 TCAGTGTTCC CAGCCTTCTC CCTGACTCTC CTGACAACTG GAGGATAAGG GAGAGTCCCC
1501 AGTCTGGTCT TCCCTCCCCA TCTCCCTACA TTGGGGCCTC TCCATGTGTA TCCCTCACCT
1561 GTCTCCAGCG GCCCTGTCCT GATTCCTCCC TGCCTCTCTC TGCCCCACCT CCTTATTCTC
1621 TCTCACTGGA GTCTCCTCTT TCCCCTCTCT CTCCATCTCT AAGGACATCC TGGGTTTCTG
1681 TTTACCAGCC CTGGGTCTCT GTCTACATGA GTCTTTGAGG CCCTCTTAGC TTCTGGGCTT
1741 CTCTGGGTTT CTCATCTCTC CGGATCCCTT TCTCAATTCT TCCTCTGTCT TAGGATGCCA
1801 GGGTTATTCC TACTTCCACA TCTTCAGGCT CCATCTCCTG GTAACAGTCT CTCTTCCTTC
1861 CAGACCCTCT CTGTTTCTAT CTCAATATTA AACTCTCTGC TCCAGCTCAG CTTAAGAATC
1921 TCACACCAAG AGAGGATGTC CTCCACCCAG ATCTCCCCAT ATCTCACTAC CCCACCCTCC
1981 ATCCTCTGCC TCCATCACTC TCTTTCTCTC CCCACTGCNC CTGCGGACGC GATCCAATGG
2041 AGTGTGGAGC TAATGCCGTG AAGCTATGTG CATCTCTCTG TCTGGCCGTA CCTGGGTAAT
2101 AACCTGATCG ACTAGGCGTG GTATTCAGCA ACGGGGAGCG CGCCAAGCAG CTCCTGCGCT
2161 TTGCCATCGC CACCCTGAGG GACTTCGGGG TGGGCAAGCG AGGCATCGAG GAGCGCATCC
2221 AGGAGGAGTC GGGCTTCCTC ATCGAGGCCA TCCGGAGCAC GCACGGTGAG CAGGGGACCC
2281 CGAGTGCGGG GGCAGGAGAA GGAAAACACC CAGGACGAGG AACCCGCGCG CGTTCTGCCT
2341 GGGGATGGGG ACTAGGTGGG GAAAGGCGCC CGCACTTCCA GCCCTGGAGT CTGGCGCTGG
2401 GAATTTGGCT CAACAAGGCC CTGCCTCCTG GAATTCTGAC TCTCCTCAGA CCTCTGAGTT
2461 GACTCTCTCC CCAACCCCCT TCTCCCGACA TACCCGGAGG CGCCAATATC GATCCCACCT
2521 TCTTCCTGAG CCGCACAGTC TCCAATGTCA TCAGCTCCAT TGTCTTTGGG GACCGCTTTG
2581 ACTATAAGGA CAAAGAGTTC CTGTCACTGT TGCGCATGAT GCTAGGAATC TTCCAGTTCA
2641 CGTCAACCTC CACGGGGCAG GTAATGGTTG CAGCCCGGCC CGTGAAGGCC CTTACCAAAA
2701 CCGGCAAATT GTTCCCCTAC CGGGGGAAGG GGGCCCCAAA TTCCCACCGC CCCCCGGACA
2761 GTGTCCCCTC AAAATCAGTC CCCGATTTGG GCAAATTGGC AGAGTGGAAC CAGACCCGGG
2821 TTGGTTGTCC AATCCCCTGC TCTCCAGGGA CACCGGGATA GCACAACAGA TGCTCCCCAA
2881 AACAGAGCCT GCTGGCAGGA TGCATACCCT CAGCTCAGCT CTCTCACCCT GGGCACGTGT
2941 TCCCATCCCC AACTTACCGG TAATTTCTAA CAGATGCTCC CTACCCAGGT CTTCTTGAAT
3001 ATTTTAACAC CCGGAAACCC TGGGTACCTA ACCTTCCCTG TAAACTTTAG AGATTAGTTC
3061 CTATCCGGCC CCTCTGAAAT ACCTAACCAC CGGAGACCAG ATGCCTTTAA CTCAGTTCCT
3121 TCCTTGCTAT GAAACAAATC CCATTCCCAT CAGCTCCTGC CCCGTGACAG CTGTCCTTCC
3181 CTTCCCATCC TCTCTCTGCA ACCCCAGCTC TATGAGATGT TCTCTTCGGT GATGAAACAC
3241 CTGCCAGGAC CGCAGCAACA GGCCTTTCAG TTGCTGCAAG GGCTGGAGGA CTTCATAGCC
3301 AAGAAGGTGG AGCACAACCA GCGCACGCTG GATCCCAATT CCCCACGGGA CTTCATTGAC
3361 TCCTTTCTCA TCCGCATGCA GGAGGTACAC CCCAGCAGCC ACTGCGGGGA GATGCAAAGC
```

FIG. 12 (Sheet 1)

3421 CAGGCAGAGG GAAATCAGTC TGGGAGTGGG GCAGGCAGAT GACACAGGCC CATTCAAATT
3481 AACCCTCATC ATAATAATCC TCACAATTGG CTGGGTGCCG TGGCTAACAG CCTGTAATCC
3541 CAGCACTTTG GGAGGCCGAG GCAGGTGGAT CACCTGAGGT CAGGAGTTCG AGACCAGCCT
3601 GGCCAACATG GTCAAACCCC GTCTCTACTA AAAATCCAAA AATTAGTTGG GCATGGTGGC
3661 GCGAAGGGGG GCAGAGGTTG CAATGAGCCA AGATCACGGC ATTGCACTCC AGTCTGGGTG
3721 ACAGAATGAG GCCCTGTGTC AAAAAAAATT AATCACTTGT TTAAAAAGTA AGTGAGCCTG
3781 CATGGTCATG CGCATGTGCA GCTCCAGCTA CTCAGGAGGC TGAGGCTGGA GGATTGCTTG
3841 AGCTCAGGAG TTGGCGTCCG GCCTGTGCAA CTTAGCAAGA CCAAGTCAGT ATAAGAAAAA
3901 AAAAAAACAA AAAAAAAGCT GACAGCTAAG TTGATAATTG ACGGACAGAT GGTCAGCAAG
3961 GTAACGAAGG TGAGAAGGAA GAGCATTGGG GGCAACGCCA GGAGTCAGGG CAAGGGCTGG
4021 TTCCTAGAGC GAGTCTGGTA GGATCTAGGG CCCCTCTTCT CCACCCTGCG GTCTTGCCCC
4081 AAAGAGAGGT CGAGGGTGCT GGGATTGCGC TAGACTCGAG TCTGTGTAGA TCTTGGGGTC
4141 CCCTCTTGAC CCCCATTGGT CTGAACCTAA GAGTGGAAGA TCCATGGGGT GAACCCCTAG
4201 ATGGTGCCCT GAGGTCAAGC AGGAGTGAGG TTGTCCTAAA GCCCCCTCTC CCTTCAGGAG
4261 GAGAAGAACC CCAACACGGA GTTCTACTTG AAGAACCTGA TGATGAGCAC GTTGAACCTC
4321 TTCATTGCAG GCACCGAGAC GGTCAGCACC ACCCTGCACT ATGGCTTCTT ACTGCTCATG
4381 AAGCACCCAG AGGTGGAGGG TAAGGCTGGA GGGGGACGGA AGTGGAGGGC CCCAGACCCT
4441 CAAAATTCCC CTTCGACTGG TGCAATGTCC CCACCTGTCC CAGATCCCGG GACCCTGAGA
4501 CGTGACTTGC TGTCCAGAGA CAGGGCAACA TTCAGCTGGT AGGCATCAGC TGAGTCTCAT
4561 TAGATATTAA AATATTGAAA ATGTCTGCAC TGATTGGTCA GTCACTTCTG TCCCAAGCCC
4621 ACTGAGTGCC CACTGCCCGT TCCACCGGGT CATCCCCTAA GTTCCTCCCT GTGCCTCCCC
4681 TGTGATTCTG GCACAACCTG GTTAACAGGA TCCTACTCCA ACAATGCGAA TGGGTGATGT
4741 CTGTTCTGTT ATGAATGCTC TACTTCCGTC TCATAGGCGG AGGCATTTCA TCCACCCCAT
4801 TTTGCCTATC CGGACTATCA TTTCCTGCTC TGAGACCCCT AGATACCTAA ACACATTCCC
4861 CCTCCTCCCC CAGCCAAGGT CCATGAGGAG ATTGACAGAG TGATCGGCAA GAACCGGCAG
4921 CCCAAGTTTG AGGACCGGGC CAAGATGCCC TACATGGAGG CAGTGATCCA CGAGATCCAA
4981 AGATTTGGAG ACGTGATCCC CATGAGTTTG GCCCGCAGAG TCAAAAAGGA CACCAAGTTT
5041 CGGGATTTCT TCCTCCCTAA GGTGCTATCC GCCCCCACCC CCCAGACTAC GGGGACTCCA
5101 GCCCCTCTCT GTGTCCCCAG CATCCCACCC ACATTAGAAG CTTTCTAGAC CCTGTCCCAC
5161 TCCCTCAATC AGTCAAAAAA GACTTCCCCA ACCACCACAT CCGTTCCACC TTTCCACTTA
5221 GACACTCCTG AGTCCTGCAT CTCTCCAGAC TCTTTGTGTC AGGAGAATCA AACACATGTT
5281 CCCAAACTTC CTATCTTAAG AAACAGAAGC CCCCTTTCCA TTCGGCCTTT TGTCATAGGG
5341 ACAGAAATCT CAGGTCCCCC AAACTCCTGC CTAGAAGGAC ATGGACCCCA TGTCTCCCAA
5401 ACTTCCTGTT TCAGAGATGT GAACCTTCTA TCCCCCAAGG TCCTCCCTCA GAGGTCCCCA
5461 ATTCCCATGC CTGCCACTTC CCCTCACCGG GGCACCCTAG TTCCCCCTCC AGCCCCTGTG
5521 TACTCTCAAC AATCCCCCAA CCCGCCTCAT CACATACACC TTCCTCCTCC CTCCCAGGGC
5581 ATAGAAGTGT TCCCTATGTT GGGCTCCGTG CTGAGAGACC TCAGGTTCTT CTCCAACCCC
5641 CGGGACTTCA ATCCCCAGCA CTTCCTGGGT GAGAAGGGGC AGTTTAAGAA GCGTGATGCT
5701 TTTGTGCCCT TCTCCATCAG TAAGAGACCA CTGTTTGGTG CCAGGCTTAC TACTCACACC
5761 AGCAGGGGCC TCCCTTACCC AGTTCCCCTC TCTGCCGTGT AGCCTAGTAT TTCCCCAGCT
5821 TGGCAAGTTC CTGTTAGCAA TCTACCGTCG AGCCACCAGG TGATACTCCC TTAACTACCA
5881 AGCACCCAGT ACCTGTGCCC AGGCAAAAGG AAAGGAAACA TCATACCCCT TTCAGAGGCG
5941 GGGGAAAACC AAAGGCCAGA GAGAATCAGA GATTTATTTC CCTAGGGTCA CACAGGAGAT
6001 TCTTCAGCAT CCCTAAAAAG GAGATGACGG CACAGCAGGT CATATTTGGG AGTTCTTATC
6061 TGGGGGAAGG GGGATCTTAA ACCTCCCATT GTGGACACCT GGCATCGATC AACCCCATCT
6121 TTTGGTCATC TTTTGGGTCA CTCAAGGAAA CTGAGGTCAA GGAGGGTCAA GAGGCTCCCT
6181 CTTAAAGTCT CTCAGGGCCA TATATTCCAC CCTTCCTCCC TGGGAGAGCC GCAGCTGGAG
6241 GTCGGTACTG GGGCGAGGCT GCACTGAGAG TGGGCTTCAC CTCCACCCCT CCCGCCTCTC
6301 CTCCTCAGGA AAGCGGAACT GTTTCGGAGA AGGCCTGGCC AGAATGGAGC TCTTTCTCTT
6361 CTTCACCACC GTCATGCAGA ACTTCCGCCT CAAGTCCTCC CAGTCACCTA AGGACATTGA
6421 CGTGTCCCCC AAACACGTGG GCTTTGCCAC GATCCCACGA AACTACACCA TGAGCTTCCT
6481 GCCCCGCTGA GCGAGGGCTG TGCCGGTGAA GGTCTGGTGG GCGGGGCCAG GGAAAGGGCA
6541 GGGCCAAGAC CGGGCTTGGG AGAGGGGCGC AGCTAAGACT GGGGGCAGGA TGGCGGAAAG
6601 GAAGGGGCGT GGTGGCTAGA GGGAAGAGAA GAAACAGAAG CGGCTCAGTT CACCTTGATA
6661 AGGTGCTTCC GAGCTGGGAT GAGAGGAAGG AAACCCTTAC ATTATGCTAT GAAGAGTAGT
6721 AATAATAGCA GCTCTTATTT CCTGAGCACG TACCCCCGTG TCACCTTTGT TCAAAAACCA
6781 TTGCACGCTC ACCTAATTTG CCACAAAACC CCCTTCGAAG GGGCGTTCAT GCCCATTTTA

FIG. 12 (Sheet 2)

```
6841 CACGTGACAA AACTGAGGCT TAGAAAGTTG TCTCTGATGT CTCACAAAAC ATAAGTGCCC
6901 AGAAAATCTG CGAACACAGA TCTGTGCCCA TAGCCTTCTA GACAGATTCT TAAAAAGCAC
6961 CTATTCCTCA CGCAAAACAG TTTAGTATAG AATCACATGG CCTGAACATC CCTGTCCGGG
7021 GGAGTTCCCC AGAGACCTGG GGGGTGGTTG CCCTGCCTTC ACTGCACACA TGCCCACACT
7081 CTCACCTACT CAACATGCTG TGACTACCCG GGTGTAATCT GTGCTTGCTA CCAGATAAGG
7141 CCACTGTAGC CCATTCAGAG TCAGCCCAGG GACACAACGA GACATGACTG GACATACAGG
7201 GTCAGTCCAT TAACAA
```

FIG. 12 (Sheet 3)

CYP2A13 cDNA

```
5'ATGGCCACCATGCTGGCCTCAGGGCTGCTTCTGGTGACCTTGCTGGCCTGCCT
GACTGTGATGGTCTTGATGTCAGTCTGGCGGCAGAGGAAGAGCAGGGGGAA
GCTGCCTCCGGGACCCACCCCATTGCCCTTCATTGGAAACTACCTCCAGCTGAA
CACAGAGCAGATGTACAACTCCCTCATGAAGATCAGTGAGCGCTATGGCCCT
GTGTTCACCATTCACTTGGGGCCCCGGCGGGTCGTGGTGCTGTGCGGACATGAT
GCCGTCAAGGAGGCTCTGGTGGACCAGGCTGAGGAGTTCAGCGGGCGAGGCGA
GCAGGCCACCTTCGACTGGCTCTTCAAAGGCTATGGCGTGGCGTTCAGCAACG
GGGAGCGCGCCAAGCAGCTCCGGCGCTTCTCCATCGCCACCCTAAGGGGTTTTG
GCGTGGGCAAGCGCGGCATCGAGGAACGCATCCAGGAGGAGGCGGGCTTCCTC
ATCGACGCCCTCCGGGGCACGCACGGCGCCAATATCGATCCCACCTTCTTCCTG
AGCCGCACAGTCTCCAATGTCATCAGCTCCATTGTCTTTGGGGACCGCTTTGA
CTATGAGGACAAAGAGTTCCTGTCACTGTTGCGCATGATGCTGGGAAGGTTC
CAGTTCACGGGAACCTCCACGGGGCAGCTCTATGAGATGTTCTCTTCGGTGAT
GAAACACCTGCCAGGACCACAGCAACAGGCCTTTAAGGAGCTGCAAGGGCT
GGAGGACTTCATCGCCAAGAAGGTGGAGCACAACCAGCGCACGCTGGATCCC
AATTCCCCACGGGACTTCATCGACTCCTTTCTCATCCGCATGCAGGAGGAGGA
GAAGAACCCCAACACAGAGTTCTACTTGAAGAACCTGGTGATGACCACCCT
GAACCTCTTCTTTGCGGGCACTGAGACCGTGAGCACCACCCTGCGCTACGGTTT
CCTGCTGCTCATGAAGCACCCAGAGGTGGAGGCCAAGGTCCATGAGGAGATT
GACAGAGTGATCGGCAAGAACCGGCAGCCCAAGTTTGAGGACCGGGCCAAG
ATGCCCTACACAGAGGCAGTGATCCACGAGATCCAAAGATTTGGAGACATG
CTCCCCATGGGTTTGGCCCACAGGGTCAACAAGGACACCAAGTTTCGGGATT
TCTTCCTCCCTAAGGGCACTGAAGTGTTCCCTATGCTGGGCTCCGAGCTGAGA
GACCCCAGGTTCTTCTCCAACCCCCAGGACTGCAGTCCCCAGCACTTCCTGGAT
GAGAAGGGGCAGTTTAAGAAGAGTGATGCTTTTGTGCCCTTTTCCATCGGA
AAGCGGTACTGTTTTGGAGAAGGCCTGGCCAGAATGGAGCTCTTTCTCTTCT
TCACCACCATCATGCAGAACTTTCGCTTCAAGTCCCCTCAGTCGCCTAAGGAT
ATCGACGTGTCCCCCAAACACGTGGGCTTTGCCACGATCCCACGAAACTACAC
CATGAGCTTCCTGCCCCGCTGAGCGAGGGCTGTGCTGGTGCAGGGCTGGTGGGC
GGGGCCAGGGAAACGGCCGGGGCAGGGGCGGGGCTTGTGGAGGGGCGGGGCT
AAGAATGGGGGCAGTGGGGGAAGGAAGGGGAGAGGTGGTTAGAGGGAACA
GAAGAAACAGAAGGGGCTCAGTTCACCTTGATGATGTCCTTCAGAGCTGTG
ATGAGAGGAAGGGAAACCTTACAGTATGCTACAAAGAGTAGTAATAATA
GCAGCTCTTATCTCCTGA 3'
```

FIG. 13

```
3421 CCTCCACTTC AGCATCTTCA CCAGCCCCAC TTTATACCTG AGCACCTGAA CAAAAGCCCC
3481 CAATCCAGAC CCAGTAAGTA TCTGGACAGC TGTCTCCAAC CAAGTCCACT TGAATGCCTA
3541 AATACCTAGA CAGGTGCCAC TCACCTCATA CCAGCCCCAC CTGAAGAGCT AAACACCTGG
3601 ACAGGTGTCT TCCAACTCAA CTTCACTTGA ATATCTGAAC ACCTAGATGT GTGCTCCAAT
3661 CCAGCCTCAT TTGCATACCT GAAACCTGGA TATATGCCTC AGTTCTTCTC ACCTAAATTA
3721 CTAGACCGTG CCCCTGGCAC CTAATCCACG TGAAAACTTA GATATAAGTT TCCATCCAAC
3781 CCCACTGAAA TACCTAAACA CCTGGACAGA TGCCTTTAAC TCCGTTCCTT CCTTGCTATG
3841 AAACAAATCC CCATTCCCAT CAGCTCCTGC CCCGTGACAG CTGTCCTTCC CTTCCCATCC
3901 TCTCTCTGCA ACCCCAGCTC TATGAGATGT TCTCTTCGGT GATGAAACAC CTGCCAGGAC
3961 CACAGCAACA GGCCTTTAAG GAGCTGCAAG GGCTGGAGGA CTTCATCGCC AAGAAGGTGG
4021 AGCACAACCA GCGCACGCTG GATCCCAATT CCCCACGGGA CTTCATCGAC TCCTTTCTCA
4081 TCCGCATGCA GGAGGTACAT CCCAGCAGCC AGTGCAGGCA GGTGCAAAGC CAGGGAGAGG
4141 GAAATCAGGA TGGGAGTGGG GTGGGCAGAC GACACAGGCC CATTCAAATT AGCCCTCGTC
4201 ATAATAATCC TTACAATTGG CCAGGCGCGG TGGCTCATGA CCTGTAATCC CAGCACTTTG
4261 GGAGGCCGAG GCAGGTGGAT CACCTGAGGT CAGGAGTTCG AGACCAGCCT GGCCAACATG
4321 GTGAAACCCC GTCTCTACTA AAAATACAAA AATGAGCTAG GTATGGTGGC ATGCGCCTGT
4381 AATCCCAGCT ACTCAGGAGG CTGAGACAGA AGAATTTGTT TGAATCCGGG AGGCAGAGGT
4441 TGCAGTGAGC CGGGATCATG CCACTGCACT CCGGCCTGAG TGACAGAGCA AGACCCTGTA
4501 AAAAAAAAAA AAAAAAAAAA AAAAAATTCC GGAAAACCCC AATTACATCA CCCACTGCTG
4561 TCCCATCTAC TGAGCCCTCA CCCACAAGGA CGGGTTATGG AGGTGGATTA GATTGGAAAG
4621 AACTTCTCAA GAACTACCGG GTGCCAGGAA CTGGGTTAAG TGTTTTATGA TAGTCCGCCA
4681 TGGAACACTT TTAACAGTTC TTGAGGGAGG TTCACTCATG GCCCCAGTTG TACAAATGAG
4741 GAAACTGAGG CCCAGAGAGT TTAAGTGTCT TAACTGAGGT CACAACAGTG AGGAAGACCA
4801 TGGTCCCCCT AGCTCAAACC CTGGTCTCTC TGAGCCTATA GCTGGTGCTT TTAGCCACCA
4861 TGCTCTCTAA CCGTTCATGT CCTGGTTAGC AGACACACCT CTGTGGACAG GTGACCTGGC
4921 TTTACATTGC AGGGTCCCCG CCTACCTCTG GATGTCAGCC TCCCATGTGG GAAGGCTTTA
4981 GGAAGCCAAA GCTCAGGGAG AAAGGATCAA GGGAGGGATT CCTCCACAGT AAGTTTCAAG
5041 ATTTTTAGGG AAGAAATAGG ATGCTGTTGC TTAAAATTCT GTGCTTGTAT CTCAGAAAAA
5101 CTCTTTTTTT CTGACTCTTC ATCTTGCCAT CTCTGTACTA CTTTCTCTTC GTCTCCCCTC
5161 ATCCTTCTCT TTCCAAATAT TCCTATCATT AAAAAAGTAA CAGACTGGGA AACATGGCAA
5221 AACCCCGTCT GTACAAAAAA ATGGCTAGGC ATGGTGGTGC ATGCCTGCGG TCCCAGCTAC
5281 TAAGGAGGTT GAGGTGGGAG GATATCTTGA GCCCAGGGTG GGCAGAGGTT TCAATGAGCC
5341 GATATCACAG CCCTGCCCTC CAGCCTGGGT GACAGAATAA GACCGTGTCT CCCAAAAAAA
5401 AAAAGAATTA ATTTTTTAAC AGTTAACAAG TGAGCCTGCA TAGTCATGTG CATGTGCAGT
5461 TCCAGCTACT CTGGAGGCTG AGACCGGAGG ATTCCTTGAA CCCAGGAGTT GGAGTCCAGC
5521 CTGTGCAACT TAGCAAGACC AAGTCTGCAT AAAAAAAAAA AAAACCAACT GACAGCTAAG
5581 TTGACAATTA AAGGATAGAT GATCAGTGAG GTAAAGAAGG TGAGAAGGAA GAGCATTTTG
5641 GGCAAAGCCA GCAGCCAGGG CAAGGGCTGG AACCTGGAGC GAGTTTGGCA AATCTAGGGT
5701 CCCTCTTTCC ACCTTTGGTC TGGACCAAAG AGAGGTAGCT CCAAAGGAAA AGCCCTAGAA
5761 GGGCCCCAAG AGCATGGAGA GTGAGCTTGG TCTAAACCGC CCTCTCCCTG CAGGAGGAGA
5821 AGAACCCCAA CACAGAGTTC TACTTGAAGA ACCTGGTGAT GACCACCCTG AACCTCTTCT
5881 TTGCGGGCAC TGAGACCGTG AGCACCACCC TGCGCTACGG TTTCCTGCTG CTCATGAAGC
5941 ACCCAGAGGT GGAGGGTAAG ACTGGAAAGG GAGGAAAGTG AAGGGCCCCA GACCCTCAAA
6001 ACTCCCCTGA GCCTGGTGCA GTGTACCCAC CTATCCCAGA TCCCAGGACC CTGAGACGTG
6061 CCTTGCTGTC CAGAGACAGG ACAATATTCA GCTGATAGGC ATCAGCTGAG TCTCATTAGC
6121 TATTAAAATA TTGAAAATGT CTGCACTGAT TGGTCAGTCA CTCCTGTCCC AAGCCCACTG
6181 AGTGTCCGCT GCCTGCTCCT CTGGATCATC CCCTAAGTTC CTCCCTTGTC CTACCCTGTG
6241 ATTCTGACAC AACCTGGTTT AACAGGGATC CTGCTGCAAA CAATGCGAAT GGGTGATGTC
6301 TTGTTCTTGT TTATGAATGG GCTTACCCTT CGTGTCAGAG GTGGAAGCTA TGTCAACCGC
6361 CGTGTTTTAG CTAGGGGGGG CGATACATGC CCTGCTCTAA GACCCCTAGA GAGGGTAAAG
6421 ATATTCCCCT CCTCCGCCAG CCAAGGTCCA TGAGGAGATT GACAGAGTGA TCGGCAAGAA
6481 CCGGCAGCCC AAGTTTGAGG ACCGGGCCAA GATGCCCTAC ACAGAGGCAG TGATCCACGA
6541 GATCCAAAGA TTTGGAGACA TGCTCCCCAT GGGTTTGGCC CACAGGGTCA ACAAGGACAC
6601 CAAGTTTCGG GATTTCTTCC TCCCTAAGGT GCTGTCTCCC CTCCACCACC ACCACTCAGA
6661 CTACGGGGAC TTCCAGCCTC TCTCTGTGTC CCCAGAATCC TGCCCCCATT AGTGTTCTAG
6721 ACTCTGTCCC ACTCCCTCAA TCAGTCAAAA AAGACTTCCC CAACCACCAC ATCTGTTCCA
6781 CCTTTCCACT TAGACAGTCC TGAGTCCTGC ATCTCGCCAG ACTCTTTGTG TCAGGAGAAT
```

FIG. 14 (Sheet 1)

```
3421 CCTCCACTTC AGCATCTTCA CCAGCCCCAC TTTATACCTG AGCACCTGAA CAAAAGCCCC
3481 CAATCCAGAC CCAGTAAGTA TCTGGACAGC TGTCTCCAAC CAAGTCCACT TGAATGCCTA
3541 AATACCTAGA CAGGTGCCAC TCACCTCATA CCAGCCCCAC CTGAAGAGCT AAACACCTGG
3601 ACAGGTGTCT TCCAACTCAA CTTCACTTGA ATATCTGAAC ACCTAGATGT GTGCTCCAAT
3661 CCAGCCTCAT TTGCATACCT GAAACCTGGA TATATGCCTC AGTTCTTCTC ACCTAAATTA
3721 CTAGACCGTG CCCCTGGCAC CTAATCCACG TGAAAACTTA GATATAAGTT TCCATCCAAC
3781 CCCACTGAAA TACCTAAACA CCTGGACAGA TGCCTTTAAC TCCGTTCCTT CCTTGCTATG
3841 AAACAAATCC CCATTCCCAT CAGCTCCTGC CCCGTGACAG CTGTCCTTCC CTTCCCATCC
3901 TCTCTCTGCA ACCCCAGCTC TATGAGATGT TCTCTTCGGT GATGAAACAC CTGCCAGGAC
3961 CACAGCAACA GGCCTTTAAG GAGCTGCAAG GGCTGGAGGA CTTCATCGCC AAGAAGGTGG
4021 AGCACAACCA GCGCACGCTG GATCCCAATT CCCCACGGGA CTTCATCGAC TCCTTTCTCA
4081 TCCGCATGCA GGAGGTACAT CCCAGCAGCC AGTGCAGGCA GGTGCAAAGC CAGGGAGAGG
4141 GAAATCAGGA TGGGAGTGGG GTGGGCAGAC GACACAGGCC CATTCAAATT AGCCCTCGTC
4201 ATAATAATCC TTACAATTGG CCAGGCGCGG TGGCTCATGA CCTGTAATCC CAGCACTTTG
4261 GGAGGCCGAG GCAGGTGGAT CACCTGAGGT CAGGAGTTCG AGACCAGCCT GGCCAACATG
4321 GTGAAACCCC GTCTCTACTA AAAATACAAA AATGAGCTAG GTATGGTGGC ATGCGCCTGT
4381 AATCCCAGCT ACTCAGGAGG CTGAGACAGA AGAATTTGTT TGAATCCGGG AGGCAGAGGT
4441 TGCAGTGAGC CGGGATCATG CCACTGCACT CCGGCCTGAG TGACAGAGCA AGACCCTGTA
4501 AAAAAAAAAA AAAAAAAAAA AAAAAATTCC GGAAAACCCC AATTACATCA CCCACTGCTG
4561 TCCCATCTAC TGAGCCCTCA CCCACAAGGA CGGGTTATGG AGGTGGATTA GATTGGAAAG
4621 AACTTCTCAA GAACTACCGG GTGCCAGGAA CTGGGTTAAG TGTTTTATGA TAGTCCGCCA
4681 TGGAACACTT TTAACAGTTC TTGAGGGAGG TTCACTCATG GCCCCAGTTG TACAAATGAG
4741 GAAACTGAGG CCCAGAGAGT TTAAGTGTCT TAACTGAGGT CACAACAGTG AGGAAGACCA
4801 TGGTCCCCCT AGCTCAAACC CTGGTCTCTC TGAGCCTATA GCTGGTGCTT TTAGCCACCA
4861 TGCTCTCTAA CCGTTCATGT CCTGGTTAGC AGACACACCT CTGTGGACAG GTGACCTGGC
4921 TTTACATTGC AGGGTCCCCG CCTACCTCTG GATGTCAGCC TCCCATGTGG GAAGGCTTTA
4981 GGAAGCCAAA GCTCAGGGAG AAAGGATCAA GGGAGGGATT CCTCCACAGT AAGTTTCAAG
5041 ATTTTTAGGG AAGAAATAGG ATGCTGTTGC TTAAAATTCT GTGCTTGTAT CTCAGAAAAA
5101 CTCTTTTTTT CTGACTCTTC ATCTTGCCAT CTCTGTACTA CTTTCTCTTC GTCTCCCCTC
5161 ATCCTTCTCT TTCCAAATAT TCCTATCATT AAAAAAGTAA CAGACTGGGA AACATGGCAA
5221 AACCCCGTCT GTACAAAAAA ATGGCTAGGC ATGGTGGTGC ATGCCTGCGG TCCCAGCTAC
5281 TAAGGAGGTT GAGGTGGGAG GATATCTTGA GCCCAGGGTG GGCAGAGGTT TCAATGAGCC
5341 GATATCACAG CCCTGCCCTC CAGCCTGGGT GACAGAATAA GACCGTGTCT CCCAAAAAAA
5401 AAAAGAATTA ATTTTTTAAC AGTTAACAAG TGAGCCTGCA TAGTCATGTG CATGTGCAGT
5461 TCCAGCTACT CTGGAGGCTG AGACCGGAGG ATTCCTTGAA CCCAGGAGTT GGAGTCCAGC
5521 CTGTGCAACT TAGCAAGACC AAGTCTGCAT AAAAAAAAAA AAAACCAACT GACAGCTAAG
5581 TTGACAATTA AAGGATAGAT GATCAGTGAG GTAAAGAAGG TGAGAAGGAA GAGCATTTTG
5641 GGCAAAGCCA GCAGCCAGGG CAAGGGCTGG AACCTGGAGC GAGTTTGGCA AATCTAGGGT
5701 CCCTCTTTCC ACCTTTGGTC TGGACCAAAG AGAGGTAGCT CCAAAGGAAA AGCCCTAGAA
5761 GGGCCCCAAG AGCATGGAGA GTGAGCTTGG TCTAAACCGC CCTCTCCCTG CAGGAGGAGA
5821 AGAACCCCAA CACAGAGTTC TACTTGAAGA ACCTGGTGAT GACCACCCTG AACCTCTTCT
5881 TTGCGGGCAC TGAGACCGTG AGCACCACCC TGCGCTACGG TTTCCTGCTG CTCATGAAGC
5941 ACCCAGAGGT GGAGGGTAAG ACTGGAAAGG GAGGAAAGTG AAGGGCCCCA GACCCTCAAA
6001 ACTCCCCTGA GCCTGGTGCA GTGTACCCAC CTATCCCAGA TCCCAGGACC CTGAGACGTG
6061 CCTTGCTGTC CAGAGACAGG ACAATATTCA GCTGATAGGC ATCAGCTGAG TCTCATTAGC
6121 TATTAAAATA TTGAAAATGT CTGCACTGAT TGGTCAGTCA CTCCTGTCCC AAGCCCACTG
6181 AGTGTCCGCT GCCTGCTCCT CTGGATCATC CCCTAAGTTC CTCCCTTGTC CTACCCTGTG
6241 ATTCTGACAC AACCTGGTTT AACAGGGATC CTGCTGCAAA CAATGCGAAT GGGTGATGTC
6301 TTGTTCTTGT TTATGAATGG GCTTACCCTT CGTGTCAGAG GTGGAAGCTA TGTCAACCGC
6361 CGTGTTTTAG CTAGGGGGGG CGATACATGC CCTGCTCTAA GACCCCTAGA GAGGGTAAAG
6421 ATATTCCCCT CCTCCGCCAG CCAAGGTCCA TGAGGAGATT GACAGAGTGA TCGGCAAGAA
6481 CCGGCAGCCC AAGTTTGAGG ACCGGGCCAA GATGCCCTAC ACAGAGGCAG TGATCCACGA
6541 GATCCAAAGA TTTGGAGACA TGCTCCCCAT GGGTTTGGCC CACAGGGTCA ACAAGGACAC
6601 CAAGTTTCGG GATTTCTTCC TCCCTAAGGT GCTGTCTCCC CTCCACCACC ACCACTCAGA
6661 CTACGGGGAC TTCCAGCCTC TCTCTGTGTC CCCAGAATCC TGCCCCCATT AGTGTTCTAG
6721 ACTCTGTCCC ACTCCCTCAA TCAGTCAAAA AAGACTTCCC CAACCACCAC ATCTGTTCCA
6781 CCTTTCCACT TAGACAGTCC TGAGTCCTGC ATCTCGCCAG ACTCTTTGTG TCAGGAGAAT
```

FIG. 14 (Sheet 2)

6841	ACACCCCATG	TTCCCAATCT	TCCTGTCTTA	AGAAACAGAA	GCCCCCTTTC	CATTAGGCCT
6901	TGTGGCTTAG	GGACACAAAT	CTCAGGTCCC	TCAAACACCC	TGGCTAGTGG	AACATGGACC
6961	CCATGTCTCC	CAAACTTCCT	GTCTCAGAGA	CATGAAACTT	CTATCCCCCA	AAGCTCCTCC
7021	CTCAGAGGTC	CCCAACTCCT	CCATGTCGTG	CCACTCCCCG	CACCTGGGGG	ACCCTAGAGC
7081	CCCCTGGAGC	CCCTGTGTAC	TTTCACCAAT	CCCCCCAACC	TGGCTCATAA	CACACACCTT
7141	CCTCCTCCCT	CCCAGGGCAC	TGAAGTGTTC	CCTATGCTGG	GCTCCGAGCT	GAGAGACCCC
7201	AGGTTCTTCT	CCAACCCCCA	GGACTGCAGT	CCCCAGCACT	TCCTGGATGA	GAAGGGGCAG
7261	TTTAAGAAGA	GTGATGCTTT	TGTGCCCTTT	TCCATCGGTA	AGAGACACTG	TTTGCTGCCA
7321	GGCCACGGCT	CACACCAGCA	GGGGCCTCTC	TCACCCACCT	CCCCTCTCTG	CGGTGTAGCC
7381	TGGTATTTCT	CCAGCTTGGA	AGTTCCTGTT	AGAATCTACC	ATTGAGCCGC	CACCAGCTGA
7441	TACTCCCTTA	ACTGCCAAGC	ACCCAATACC	TGCGCCCAGG	TAAAAGGGAA	GGAAACATCT
7501	TCCCCCATAG	ATTTATTTGT	CTAGGGTCAC	ACAGCAGATT	CTTCAGCTCC	CTGAAAAGGA
7561	GATAATGGTA	CAGCACAGCA	GTCATATTTG	CAAGTGTATC	TGGGGGGTAG	GGGCATCTAA
7621	ACCTCCCATT	GCTACACCTG	GCATGGATCA	CCCCATCTAT	GATGGAGGCA	TGACATTATG
7681	CCTTTTTCGA	AACCCATAGA	ACTGTATAAC	ACAGAGTAAA	CCCTAATGTA	AACTATGGAC
7741	TTTGGTTAGT	AATAATATAT	CAATATTGGT	TCACCATTGT	TATATCTCTT	ATAGAAGGAA
7801	ACTGAAGCTC	AGGGAGGATC	GGAGTCTCCT	CTGAAAGTCT	CTCAGGCCAT	AATATTCCCA
7861	CCCCTCCTCC	CTAGAGAGTG	CAGCCGGGGG	TCAGTAGGGG	TTGAGGCTGC	ACTGAGAGTG
7921	GGCTTCACCT	TCACCCCTCC	TGCCTCTCCT	CCTCAGGAAA	GCGGTACTGT	TTTGGAGAAG
7981	GCCTGGCCAG	AATGGAGCTC	TTTCTCTTCT	TCACCACCAT	CATGCAGAAC	TTTCGCTTCA
8041	AGTCCCCTCA	GTCGCCTAAG	GATATCGACG	TGTCCCCCAA	ACACGTGGGC	TTTGCCACGA
8101	TCCCACGAAA	CTACACCATG	AGCTTCCTGC	CCCGCTGAGC	GAGGGCTGTG	CTGGTGCAGG
8161	GCTGGTGGGC	GGGGCCAGGG	AAACGGCCGG	GGCAGGGGCG	GGGCTTGTGG	GAGGGGCGGG
8221	GCTAAGAATG	GGGGCAGTGG	GGGAAGGAAG	GGGAGAGGTG	GTTAGAGGGA	ACAGAAGAAA
8281	CAGAAGGGGC	TCAGTTCACC	TTGATGATGT	CCTTCAGAGC	TGTGATGAGA	GGAAGGGAAA
8341	CCTTACAGTA	TGCTACAAAG	AGTAGTAATA	ATAGCAGCTC	TTATCTCCTG	AACAAGTCCC
8401	TCCCTGTCAG	CTTTGTTCAA	AAAGCGTTGC	ACGCTCACCT	CACTTATTTG	CCACACACCT
8461	CTACCAATGG	GGGAAAAGTC	TTCATTCCCC	TTTTTACACG	TGAGAAAGGT	GCGGCTCAGA
8521	AAGTTGTCTC	TATCTGAAAA	CTCACAAAAC	GCAAGTGTCC	AGAGGATCTT	GGAACACAGA
8581	TCTGGGCCCA	TAGCCCTCTA	GATCGATCCT	CACCATAGCA	CCCCTTCTTC	ACGTAAAATA
8641	GCTTAGTATA	GCATCACATG	GCCTGAACAC	CCCTGGGCCG	GGGGGTTCCC	CAGAGACCTG
8701	GCGGGCGGCT	GCCCTGCCTA	CTCTGTACAC	TCGCCTACTC	GGGACGATCC	GGGCACCAGG
8761	GTGTCACCTG	AGCTCGCTA				

FIG. 14 (Sheet 3)

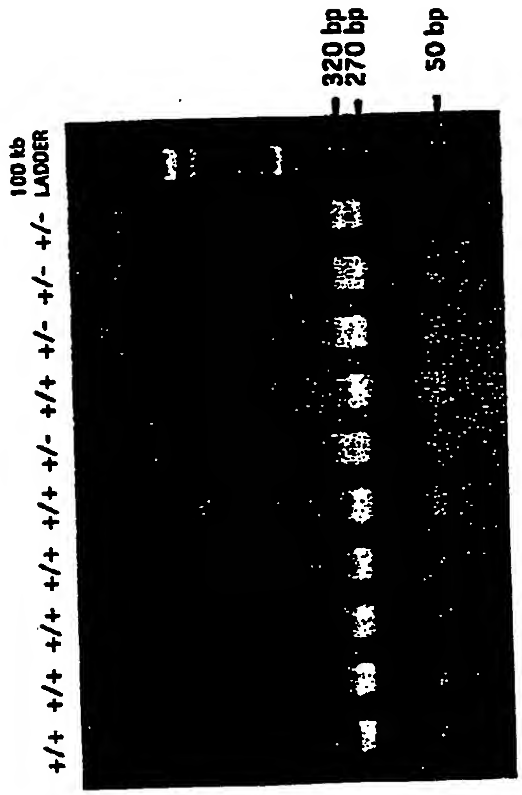


FIG. 15